

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 6, 2004, 18:18:04 ; Search time 833 Seconds

(without alignments)  
4584.790 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGMLPALLGIVLAWL.....AVAGLQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 segs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USFTO.spool/US09970944/runat\_05102004\_112009\_1616/app\_query.fasta\_1.1095  
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09970944 -CGN\_1\_1\_581 -runat\_05102004\_112009\_1616 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4787	100.0	2881	6	Abk49422 DNA encod
2	4698.5	98.2	2752	6	Abk37922 cDNA enco
3	4688.5	97.9	2697	6	Abk52891 Human net
4	4545.5	95.0	2697	6	Aas16843 Rat retri
5	4545.5	95.0	3014	2	Aav52940 Rat UNC-5
6	4434	92.6	2907	4	Aak52261 Human pol
7	4379.5	91.5	3580	6	Abk15169 Human REP
8	2694	56.3	1787	2	Aav52941 Human UNC

9	2673.5	55.8	3646	5	AAS75738	DNA encod
10	2490.5	52.0	2935	6	ABK32062	DNA encod
11	2484	51.9	2860	6	ABT06279	Human NOV
12	2484	51.9	2860	6	ABT06280	Human NOV
13	2479	51.8	3884	4	AAS21316	Human cDN
14	2479	51.8	3884	7	ACD23925	Novel hum
15	2479	51.8	3884	7	ACA67066	cDNA enco
16	2479	51.8	3884	7	ACA03675	cDNA enco
17	2479	51.8	3884	7	ABX89213	DNA encod
18	2479	51.8	3884	7	ACD41867	Human sec
19	2479	51.8	3884	7	ACA04096	Human cDN
20	2479	51.8	3884	8	ADA45664	Novel hum
21	2479	51.8	3884	8	ADA76095	Human PRO
22	2479	51.8	3884	8	ADA18745	Human PRO
23	2479	51.8	3884	8	ADA61368	Homo sapi
24	2479	51.8	3884	8	ADB19153	Novel hum
25	2479	51.8	3884	8	ADB27694	cDNA enco
26	2479	51.8	3884	8	ADA86173	Novel hum
27	2479	51.8	3884	8	ADB15737	Human PRO
28	2479	51.8	3884	8	ADA47523	Human PRO
29	2479	51.8	3884	8	ADA67318	Human PRO
30	2479	51.8	3884	8	ADB30325	cDNA enco
31	2479	51.8	3884	8	ADA85621	Novel hum
32	2479	51.8	3884	8	ADA96833	Human PRO
33	2479	51.8	3884	8	ADA79137	Human PRO
34	2479	51.8	3884	8	ADA87276	Novel hum
35	2479	51.8	3884	8	ADB16478	Human PRO
36	2479	51.8	3884	8	ADA91570	Novel hum
37	2479	51.8	3884	8	ADB14633	Human PRO
38	2479	51.8	3884	8	ADB18594	Novel hum
39	2479	51.8	3884	8	ADA93809	Human PRO
40	2479	51.8	3884	8	ADB19705	Novel hum
41	2479	51.8	3884	8	ADB13017	Human PRO
42	2479	51.8	3884	8	ACD98496	Novel hum
43	2479	51.8	3884	8	ADA74271	Human PRO
44	2479	51.8	3884	8	ADB24504	Human PRO
45	2479	51.8	3884	8	ADA82028	Human PRO

# ALIGNMENTS

RESULT 1

ABK49422

ID ABK49422 standard; DNA; 2881 BP.

XX

AC ABK49422;

XX

DT 15-JUL-2002 (first entry)

XX

DE DNA encoding human UNC5-like protein NOV1.

XX

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;  
KW cell signal processing; metabolic pathway modulation; cancerous tissue;  
KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;  
KW chromosome 13; gene; ds.

XX Homo sapiens.

XX

FH Key Location/Qualifiers

CDS 87..2786

FT /\*tag= a

FT /product= "Human UNC5-like protein NOV1"

XX

PN WO200229038-A2.

XX

PD 11-APR-2002.

XX

PF 04-OCT-2001; 2001WO-US031377.

XX

PR 04-OCT-2000; 2000US-0237862P.

XX

PA (CURA-) CURAGEN CORP.

XX



QY	561	CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly	580
DB	1767	TGTGGACCCCTGGCGTCTGCTCACCGGCCAGTCATCTCTGGTATGGACCACTGTGGG	1826
QY	581	GluProSerProAspSerTyrSerLeuArgLeuLeuLysGlnSerCysGluGlySerTyr	600
DB	1827	GAGCCACGCGCTGACAGCTGCGCTCAAAAGCAGTCGTGCGAGGGCAGCTGG	1886
QY	601	GluClnAspValLeuHisLeuGlyGluClnAlaProSerHisLeuTyrTyrCysGlnLeu	620
DB	1887	GAGCAGGATGTGTGCACCTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCCAGCTG	1946
QY	621	GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu	640
DB	1947	GAGGCCAGTGCCTGCTACGCTTCACCGAGCAGCTGGCGCGCTTGGCCCTGGTGGGAGAG	2006
QY	641	AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys	660
DB	2007	GCCTTCAGCGTGGCTGCGCGCAAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCGCTGC	2066
QY	661	ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys	680
DB	2067	ACCTCCCTCAGGTACCAATCCGGGTCTACTGCTGCATGACACCCACGATGACACTCAAG	2126
QY	681	GluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeu	700
DB	2127	GAGGTGTCAGCTGGAGAAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTTG	2186
QY	701	HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu	720
DB	2187	CACCTTCAAGGACAGTTACCAACCTCGCGCTATCCATCCACGATGTGCCCAGCTCCCTG	2246
QY	721	TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly	740
DB	2247	TGGAAGAGTAAGCTCCTTGTTCAGTACCAGAGAGTCCCTTTTATCACTCTGGATGGC	2306
QY	741	ThrGluArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp	760
DB	2307	ACGCAGCGGTACTTGCACTGCACCTTCACCGCTGGAGCGTGTACGCCCCAGCACTAGTGAC	2366
QY	761	LeuAlaCysLysLeuTrpValTrpGlnValClnGlyAspGlyGlnSerPheSerIleAsn	780
DB	2367	CTGGCTTGCAGCTGTGGGTGTGGAGGTGGAGGGCGACGGGCGAGAGCTTCAGCATCAAC	2426
QY	781	PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly	800
DB	2427	TTCAACATCACCAAGGACACAAAGTTTGTGAGCTGTGCTCTCGAGAGTGGAGCGGG	2486
QY	801	ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle	820
DB	2487	GTCCCCAGCCCTGGTGGGCCCGAGTGCCTCAAGATGCCCTTCCTCATTCGGCAGAAAGATA	2546
QY	821	IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys	840
DB	2547	ATTTCAGCTTGACCCACCTGTAGCGGGGTGCCGACATGGCGGACTCTGGGCCCAAGAA	2606
QY	841	LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle	860
DB	2607	CTCCACTGTGACAGCCATCTCAGCTTCTTTGCTCCAAAGCCCAAGCCCAAGCCATGATC	2666
QY	861	LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	880
DB	2667	CTCAACTGTGGGAGGGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGGCGCA	2726
QY	881	ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899
DB	2727	GTGGCTGGACTGGGCCAGCCAGACGCTGGCTCTTTCAGTGTGGAGGCTGAGTGC	2783
RESULT 2			
ABK37922			
ID ABK37922 standard; cDNA; 2752 BP.			
XX			

RESULT 2  
ABK37922  
ID ABK  
XX

AC	ABK37922;
XX	
DT	21-MAY-2002 (first entry)
XX	
DE	cDNA encoding Human protein NOV1.
XX	
XX	Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;
KW	cell signal processing disorder; metabolic disorder; obesity; infection;
KW	anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW	Alzheimer's disease; Parkinson's disease; immune disorder;
KW	haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW	osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW	myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW	psychosis; neurological disorder; anxiety; schizophrenia;
KW	manic depression; dementia; dyskinesia; Huntington's disease;
XX	Gilles de la Tourette's syndrome; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	WO200210216-A2.
XX	
PD	07-FEB-2002.
XX	
PF	30-JUL-2001; 2001WO-US024225.
XX	
PR	28-JUL-2000; 2000US-0221409P.
PR	04-AUG-2000; 2000US-0222840P.
PR	04-AUG-2000; 2000US-0223752P.
PR	04-AUG-2000; 2000US-0223762P.
PR	04-AUG-2000; 2000US-0223769P.
PR	04-AUG-2000; 2000US-0223770P.
PR	14-AUG-2000; 2000US-0225146P.
PR	15-AUG-2000; 2000US-0225392P.
PR	15-AUG-2000; 2000US-0225470P.
PR	16-AUG-2000; 2000US-0225697P.
PR	01-FEB-2001; 2001US-0263682P.
PR	05-APR-2001; 2001US-0281645P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;
PI	Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;
XX	
DR	WPI; 2002-180074/23.
DR	P-PSDB; AAU85403.
XX	
PT	New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT	polypeptide, useful for treating cardiomyopathy, atherosclerosis,
PT	infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT	immune disorders.
XX	
PS	Claim 9; Page 9-10; 213pp; English.
XX	
CC	The invention relates to an isolated cytoplasmic, nuclear, membrane
CC	bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature
CC	form. Also included are the nucleic acids encoding the NOVX proteins, a
CC	vector comprising the nucleic acid, a cell comprising the vector, an anti
CC	-NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC	antibody are useful for treating or preventing a NOVX-associated
CC	disorder, where the disorder is selected from cardiomyopathy,
CC	atherosclerosis, diabetes, a disorder related to cell signal processing
CC	and metabolic pathway modulation, metabolic disorders, obesity,
CC	infectious disease, anorexia, cancer-associated cachexia, cancer,
CC	neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC	immune disorders, haematopoietic disorders, and the various
CC	dyslipidaemias, metabolic disturbances associated with obesity, the
CC	metabolic syndrome X and wasting disorders associated with chronic
CC	diseases, bacterial, fungal, protozoal and viral infections, pain,
CC	bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's
CC	disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina
CC	pectoris, myocardial infarction, ulcer, allergy, benign prostatic
CC	hypertrophy, and psychotic and neurological disorders, including anxiety,
CC	schizophrenia, manic depression, delirium, dementia, and dyskinesias.





```
Db 1960 GAGGCCCTCAGCGTGGCTGCGCCCAAGCGCCTCAAGCTGCTTCTGTTTGGCCGGTGGCC 2019
Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 2020 TGCACCTCCCTCGAGTCAACATCCGGTCTACTGCTGTCATGACCCACGATGCACTC 2079
Qy 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuGlnGluProArgVal 699
Db 2080 AAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCAGGGTC 2139
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2140 CTGCACCTCAAGAGCAGTATCCACAACCTGCGCCTATCCATCCAGATGTCGCCAGCTCC 2199
Qy 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2200 CTGTGGAGAGTAGCTCTTGTGAGTACGAGAGATCCCTTTTATCATCTTGAAT 2259
Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2260 GGCACGCGAGCGGTACTTGCACTGCACCTTCACCTCGAGCGGTGCAGCCCCAGCACTAGT 2319
Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2320 GACCTGGCCTGCAAGCTGTGGTGTGGCAGTGGAGGGCGACGGGCAGAGCTTCAGCATC 2379
Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAla 799
Db 2380 AACTTCACATCACCAGGACACNAGTTTGTGAGCTGTGGCTCTGGAGAGTGAAGCG 2439
Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2440 GGGGTCCAGCCCTGGTGGGCCCGCCAGTGCCTCAAGATCCCTTCCTCATTCGGCAGAAG 2499
Qy 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2500 ATAATTTCCAGCTGGAGCCACCCCTGTGAGCGGGGTGCGGAGTCTGGCCCGAG 2559
Qy 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2560 AACTCCACTGGACAGCCATCTCAGCTTCTTTCCTCCAAAGCCCGCCAGCCCATG 2619
Qy 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db 2620 ATCTCAACTGTGGAGGGCGGCGCACTTCCCAACGGCACTCAGCCAGCTGGCTGCA 2679
Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2680 GCAGTGGCTGGACTGGGCCAGCCAGCGTGGCTCTTTCACAGTGTGGAGGTGAGTGC 2739

RESULT 3
ABK52891
ID ABK52891 standard; DNA; 2697 BP.
XX
AC ABK52891;
XX
27-AUG-2002 (first entry)
DE Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
XX
KW Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
KW neurotropic; neuroprotective; cytosolic; antiparkinsonian;
KW cerebroprotective; cancer; central nervous system; CNS; stroke;
KW Parkinson's disease; multiple sclerosis; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..2697
FT CDS /*tag= a
FT /product= "Netrin binding membrane receptor UNC5H-1"
XX
PN W0200233080-A2.
```

```
XX 25-APR-2002.
PD
XX 15-OCT-2001; 2001WO-EP011891.
PF
XX 16-OCT-2000; 2000US-0240061P.
PR
XX (FARB ) BAYER AG.
PA
XX Koehler RH;
PI
XX WPI; 2002-463314/49.
DR
XX P-PSDB; AAU97899.
DR
XX
XX Novel human netrin binding membrane receptor polypeptide and
XX polynucleotides for identifying modulating agents useful in treating
XX diseases e.g. Parkinson's disease, multiple sclerosis, stroke,
XX Alzheimer's disease.
XX
XX Claim 1; Fig 1; 94pp; English.
XX
XX This invention relates to the DNA and protein sequences of a novel
XX purified human netrin binding membrane receptor, UNC5H-1. The DNA
XX sequence of the invention is useful as a probe for detecting a nucleic
XX acid encoding the UNC5H-1 protein in a biological sample. The sequences
XX of the invention are useful to screen for agents which decrease the
XX activity of the UNC5H-1 protein. The sequences are also useful for
XX screening agents which regulate (modulate) the activity of the protein of
XX the invention. A pharmaceutical composition containing the protein of the
XX invention or a reagent that modulates the activity of the UNC5H-1 protein
XX may be useful for treating a UNC5H-1 dysfunction related disease such as
XX cancer or a central nervous system (CNS) disorders (e.g. Parkinson's
XX disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion
XX proteins comprising the UNC5H-1 protein are useful for generating
XX antibodies and for in various assay systems, and the protein can be used
XX as a bait protein in a two-hybrid assay or three-hybrid assay. The method
XX of the invention is useful for detecting a coding sequence for the UNC5H-
XX 1 protein. The present sequence represents a DNA sequence encoding the
XX human netrin binding membrane receptor UNC5H-1 protein of the invention
XX
XX Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;
```

## Alignment Scores:

Pred. No.:	4.4e-314	Length:	2697
Score:	4688.50	Matches:	886
Percent Similarity:	98.78%	Conservative:	3
Best Local Similarity:	98.44%	Mismatches:	8
Query Match:	97.94%	Indels:	3
DB:	6	Gaps:	3

US-09-970-944-2 (1-899) x ABK52891 (1-2697)

Qy	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAAlaTrpLeu	20
Db	1	ATGGCGTCCGGCCCGCTGTGGCCAGCGCTCTGGGCATAGTCTCGCCGCTTGGCTC	60
Qy	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
Db	61	CGCGGCTCGGGTGGCCAGCAGAGTGCACCGTGGCCCAACCCAGTGCCTGGTGCACACCG	120
Qy	41	AspLeuLeuProHisPheLeuValGluProGluAspValTrileValIleAsnLysPro	60
Db	121	GGACTCTCTCCACATTCCTGTGGAGCCCGAGGATGTGTACATCTCTCAAGAACAGCCA	180
Qy	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	181	GTGCTGCTTGTGTGCAAGCGCTGCCCGCCAGCAGATCTTCTCAAGTGCACCGGGAG	240
Qy	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	241	TGGTGGCGCAGGTGGACCACTGATCGAGCGCAGCAGACGCGGAGCAGTGGGTGCC	300
Qy	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120

Db 301 ACCATGAGGTCCGCTAATATGCTCAAGGACGAGTCCAGAAAGGTTTCGGGCTGGAG 360  
Qy 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140  
Db 361 GAATACTGGTGCCAGTCCGTGGCATGAGCTCTCCGGGCACCAAGAGTCAAGAGGCC 420  
Qy 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnProLeuAlaLysGluVal 160  
Db 421 TACATCCGCATAGCTATTTCGCGAAGAACTTCGAGCAGAGCGCTGGCCAAAGAGGTG 480  
Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180  
Db 481 TCCCTGGAGCAGGCGATCGTGTCCCTGCGCTCCACCGAGGCGCATCCCTCCAGCCGAG 540  
Qy 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
Db 541 GTGGAGTGGTCCGGAACGAGGACCTGGTGACCCGCTCCCTGGAGCCCAATGTATACATC 600  
Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
Db 601 ACOCGGGACACACCTGGTGGTGGACAGGCCGCTTGTGTGACAGCGCCAACTACACC 660  
Qy 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr 240  
Db 661 TGGTGGCCCAAGAACAATCGTGGCACTGCGCGAGCGCTCCGTGTGTCTCATCTGTAC 720  
Qy 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
Db 721 GTGAACGGTGGTGGTGGACGCTGGACAGTGGTCCGTCTGACGGCGCAGCTGTGGGCGC 780  
Qy 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
Db 781 GGCTGGCAGAAACAGGACCGGAGTGCACCAACCGCGCGCTCTCAACGGGGCGCTTTC 840  
Qy 281 CysGluGlyGlnAsnValHisAspArgValSerSerLeuValSerValAspGly 300  
Db 841 TGTGAGGGGAGATGTC---CAGAAACACCTTGGCCACCTTGGCCAGTGGACGGC 897  
Qy 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
Db 898 AGCTGGAGCCGTGGACAGTGGTGGCCCTGTGGGCTGGACTGCACCACTGGCGGAGC 957  
Qy 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
Db 958 CGTGAGTGCTGTGACCCAGACCCCGCAACGGAGGGAGGAGTCCAGGGCAGCTGACCTG 1017  
Qy 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
Db 1018 GACACCCGCAACTGTACAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 1077  
Qy 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
Db 1078 GCCCTCTATGGGCGCTCATCGCGTGGCCGCTCTGCCTGGTCTGCTGCTGTGCTC 1137  
Qy 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400  
Db 1138 ATCTCTGTTATTGCCGGAAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1197  
Qy 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
Db 1198 CTCACCTCAGGCTCCAGCCCGTCCAGCATCAAGCCAGCAAGCAGACCAACCCCATCTG 1257  
Qy 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlnSerLeuCysPro 439  
Db 1258 CTCACATCCAGCCGAGCTCAGACCAACCAACCACTACCAAGGGCAGTCTCTGTCCC 1317  
Qy 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
Db 1318 CGCGAGGATGGGCCAGGCCCAAGTTCAGTCCAGTCCACCAATGGGCACCTGCTCAGCCCTG 1377  
Qy 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479

Db 1378 GGTGGGGCGCGCACACACTGACACAGCTCTCCACCTCTCCAGGCGCAGGAGTTTCGTC 1437  
Qy 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
Db 1438 TCCCGCTCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGGAGGCACCAACATGACC 1497  
Qy 500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
Db 1498 TATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACAGGAATCAGCCTC 1557  
Qy 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539  
Db 1558 CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATCAGATCTACCTCAGCTGCAC 1617  
Qy 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
Db 1618 AAGCCGGAACGCTGAGTTGCCCTAGCTGGCTGTGAGACCCCTGCTGAGTCCCATCGTT 1677  
Qy 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
Db 1678 AGCTGTGGACCCCTGGCGTCTGTCTACCCGGCCAGTCTATCTGGCTATGGACCACTGT 1737  
Qy 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599  
Db 1738 GGGGAGCCAGCCCTGACAGCTGGAGCTGGCGCTCAAAAAGCAGTCTGTCGAGGGCAGC 1797  
Qy 600 TrpGluAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619  
Db 1798 TGGGAG---GATGTGTGCACCTGGCGAGGAGCGCCCTCCACCTTCTACTCTGCAG 1854  
Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
Db 1855 CTGGAGCCAGTGGCTGCTAGCTTTCACCGAGCAGCTGGGCGCTTTCGCTTGGTGGGA 1914  
Qy 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659  
Db 1915 GAGGCCCTCAGCGTGGCTGGCGCCAGCGCTCAAGCTGCTTCTGTTGGCCCGCTGGCC 1974  
Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
Db 1975 TGCACCTCCCTCGAGTACAACTCCGGGTCTACTGCTGTGATGACACCCAGATGCATC 2034  
Qy 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699  
Db 2035 AAGGAGTGGTGGAGTGGAGACAGCTGGGGGAGCAGCTGATCCAGGACCCAGCGGTC 2094  
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
Db 2095 CTGCACCTCAAGGACAGTTACCAACCTGGCGCTATCCATCCAGATGTGCCAGCTCC 2154  
Qy 720 LeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
Db 2155 CTGTGGAGAGTAAAGTCTTGTGCTGCTACAGGAGATCCCTTTTATCATCATCTGGAT 2214  
Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
Db 2215 GGCACGACGCGTACTTGCATGCACTTCCCTCGAGCGGTGTGAGCCCGCAGCAGTGT 2274  
Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779  
Db 2275 GACCTGGCGCTCAAGCTGTGGGTGTGGCAGTGGAGGGCGCAGCGGCGAGCTTCAGCATC 2334  
Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
Db 2335 AACTTCAACATCACCAAGGACCAAGGTTTGTGCTGCTGAGCTGTGCTGTGGAGGTGAAGCG 2394  
Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
Db 2395 GGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAG 2454  
Qy 820 IleIleSerSerLeuAspProProCysArgGlyAlaAspTrpArgThrLeuAlaGln 839  
Db 2455 ATAATTTCCAGCTGGACCCACCCCTGTAGGCGGGTGCCTGCTGCGGACTCTGGCCCGAG 2514

QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859  
 Db 2515 AAACCTCCACCTGGACAGCATCTCAGCTTCTTTGGCTCCAAAGCCAGCCGCCACAGCCATG 2574  
 QY 860 IleLeuAsnLeuTrpGluAlaAraGHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879  
 Db 2575 ATCTCAACCTGTGGAGGCGCGCACTTCCCAAGCGCAACCTCAGCAGCTGGCTGCA 2634  
 QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
 Db 2635 GCAGTGGCTGACTGGGCGCCAGCAGCTGGCTCTTCACAGTGTGGAGGCTGAGTGC 2694

RESULT 4  
 AAS16843  
 ID AAS16843 standard; cDNA; 2697 BP.  
 XX  
 AC AAS16843;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Rat netrin receptor UNC5H1 (YSG7) cDNA.  
 XX  
 KW YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;  
 KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;  
 KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;  
 KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic; ss;  
 KW tumour necrosis factor alpha; TNF-alpha; rat.  
 XX  
 OS Rattus sp.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 1..2697  
 FT /\*tag= a  
 FT /product= "Rat netrin receptor UNC5H1"  
 XX  
 FN WO200175440-A2.  
 XX  
 XX 11-OCT-2001..  
 XX  
 XX 02-APR-2001; 2001WO-GH001486.  
 XX  
 FR 31-MAR-2000; 2000GB-00007880.  
 PR 26-MAY-2000; 2000GB-00012768.  
 XX  
 XX (WELF-) WELFIDE CORP.  
 PA  
 XX  
 XX Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;  
 PI  
 XX  
 DR WPI; 2002-010813/01.  
 DR P-PSDB; AAU10543.  
 XX  
 PT Novel chronic animal model of schizophrenia, useful for identifying anti-  
 PT psychotic drugs and genes that are associated with schizophrenia.  
 XX  
 PS Claim 1; Fig 8a; 79pp; English.  
 XX  
 CC The invention relates to YSG polynucleotide fragments for use in  
 CC diagnosing and/or developing treatments for schizophrenia using chronic  
 CC animal models. The polynucleotides and their encoded polypeptides are  
 CC used for identification of compounds which modulate the expression of YSG  
 CC molecules, leading to the manufacture of schizophrenia medications. The  
 CC sequences can also be used for testing candidate compounds for any effect  
 CC on the polypeptides. Anti-schizophrenic effects of a compound can be  
 CC determined by measuring local cerebral glucose utilisation (LCGU) or  
 CC comparing its expression level with that of a control group. The  
 CC sequences are useful in the identification of genes associated with  
 CC schizophrenic states and in the development of an antibody. The sequences  
 CC of the invention include phosphodiesterase 1-alpha, calcium-independent  
 CC alpha-latrotoxin receptors (CIRL)-1,2k3, epithelial discoidin domain  
 CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and 1B and  
 CC tumour necrosis factor (TNF) alpha. This sequence represents rat netrin  
 CC receptor UNC5H1 (YSG7) DNA

XX  
 SQ Sequence 2697 BP; 541 A; 864 C; 766 G; 526 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3.32e-304 Length: 2697  
 Score: 4545.50 Matches: 852  
 Percent Similarity: 96.78% Conservative: 19  
 Best Local Similarity: 94.67% Mismatches: 26  
 Query Match: 94.96% Indels: 3  
 DB: 6 Gaps: 3  
 US-09-970-944-2 (1-899) x AAS16843 (1-2697)  
 QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaAlaTriPleu 20  
 Db 1 ATGGCCGTCGGCCCGCCGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCCGCTGGCTT 60  
 QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
 Db 61 CGTGGTTCGGGTGCCAGCAGAGTGCACCGTGGCCCAATCCAGTGGCCGGTGCCAAACCCC 120  
 QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60  
 Db 121 GACCTGCTGCCCATCTTCTGTAGAGCTGTAGAGACGTGTACATTGTCAAGAACAAAGCCG 180  
 QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
 Db 181 GTGTGTGTGGTGTGCAAGGCTGTCCCTGCCACCCAGATCTTCTTCAAGTGCATGGGAA 240  
 QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
 Db 241 TGGGTCCGCCAGGTGATCATGTAATGAACGAGCAGCCAGCAGCAGCGGATTTGCCA 300  
 QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
 Db 301 ACCATGGAGTCCGTATCAACGTATCGAGCAGCAGGTAGAGAAAGTGTGGCTGGAG 360  
 QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140  
 Db 361 GAATACTGGTGGCAGTGTGTGGCATGGAGCTCTCGGTACCAACCAAAAGTCAGAGGCC 420  
 QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGluGlnProLeuAlaLysGluVal 160  
 Db 421 TACATCGGATTGCCATTTCGCAAGAACTTTGAGCAGGAGCCACTTGGCCAAAGGAAGT 480  
 QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180  
 Db 481 TCACTGGAGCAAGGCATTGTACTACCTTGTGCCCCCAGAGAGAAATCCCCCAGCTGAG 540  
 QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
 Db 541 GTGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTATCATC 600  
 QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
 Db 601 ACCTGGAGCAGCAGCCCTAGTGTGGTCCAGCCGCTGGCCGACACGCGCAACTACACC 660  
 QY 221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaValIleValTyr 240  
 Db 661 TGTGTGCCCAAGAACATCGTACCCGTCGCCGAGACACCTTCGACGGGTCAATGTTTAT 720  
 QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
 Db 721 GTAAACGTTGGTGGTTCGACGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780  
 QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
 Db 781 GCTGTGCAGAAACGAGCGCGAGCTGCACCAACCCGCGACCTCTCAACGCGGGCGGCTTC 840  
 QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
 Db 841 TGTGAGGGGCAAGTGC---CAGAAACAGCGCTGGCCACCTCTGTGCCAGTGGATGGG 897

301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
898 AGCTGGAGTTCTGGTGAATAGTGGTCAAGCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957  
321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340  
958 CGGAGTGCTCTGACCCAGACACCCCGAATGGAGGTGAGAGGTGCGGGGTGTGACCTG 1017  
341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
1018 GACACCCGCAACTGTATCAGTGACCTCTGCTGCACACCGCTTCTTCCGCCGAGGACGTG 1077  
361 AlaLeuTyrValGlyLeuLeuAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
1078 GCTCTCTACATCGCGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1137  
381 IleLeuValTyrCysArgLysLysGlyGlyLeuAspSerAspValAlaAspSerSerIle 400  
1138 GGACTCATTTACTGTCCAGAGAGGAGGCTGGACTCCGATGTGGCGGACTGCTGCATC 1197  
401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
1198 CTCACCTCGGCTTCCAGCCTGTGAGCATCAAGCCAGCAAGCAGACAAACCCACACCTG 1257  
421 LeuThrIleGlnProAspLeuSer--ThrThrThrThrTyrGlnGlySerLeuCysPro 439  
1258 CTCACCATTCAGCCAGACCTCAGCACCACTACCACTACCAAGGAGGCTGTATGTGCG 1317  
440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
1318 AGCAGAGATGGACCCAGCCCAAGTTCAGCTCTTAATGTGTCACTGCTCAGCCACTG 1377  
460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479  
1378 GGGAGTGGCGCCATACGTTGCACACACAGCTCACCACCTCTGAGGCTGAGGACTTCGTC 1437  
480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
1438 TCCCGCTCTCCACCCCAAACTACTTTGTTTCCCTGCCCCCGGACCCAGCAACATGGCC 1497  
500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
1498 TAGGGACCTTCACTCTCTGGGGGCGGCTGATGATCCCTAATACGGGATCAGCCTC 1557  
520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539  
1558 CTCATACCCCGGATGCCATCCCCGAGGAAGATCTACGAGATCTACCTCACACTGCGAC 1617  
540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
1618 AAGCCAGAGACGTGAGGTTGCCCTAGCTGGCTGTGAGACCCCTGCTGAGTCCAGTCTGTT 1677  
560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
1678 AGCTGTGGCCCCCAGGAGTCTCTGCTCCAGCGCCAGTCACTCTTGCATAGGACCACTGT 1737  
580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599  
1738 GGAGAGCCACCCCTGACAGCTGGAGTCTGGCCCTCAAAAAGCAGATCTCGCGGGGAGT 1797  
600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619  
1798 TGGAG--GAITGCTGCACCTTGGTGAGGAGTACCTTCCACCTCTACTACTGCGCAG 1854  
620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
1855 CTGGAGGCGGGGCTGTGTGTCTTCCAGGAGAGTGGCGGCTTGGCCCTGTGTAGGA 1914  
640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659  
1915 GAGGCGCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCTCTGTTGCTCCCGTGGCC 1974  
660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679

1775 TGTAAGTCCCTTGGTGTACATCCGAGTGTACTGCTACACGACACCCACGAGCTCTC 2034  
680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuLeuGlnLeuProArgVal 699  
2035 AAGGAGTGTGTGAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGGCTC 2094  
700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
2095 CTGCACCTTCAAGACAGTGTACCAACACCTACGCTCTCCATCCACGAGTGGCCAGTCC 2154  
720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
2155 CTGTGGNAGCAAGCTACTTGTGAGTACAGGAGATCCCTTTTACCATCATCTGAAC 2214  
740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
2215 GGCACCCAGCATATCTGCACCTTCACCTTGGAGGCGCATCAACGCGACACCCAGC 2274  
760 AspLeuAlaCysLysLeuTrpValTyrGlnValGluGlyAspGlyGlnSerPheSerIle 779  
2275 GACTGSCCTTGCAGAGTGTGGTGTGSCAGGTGGAGGAGATGGGACAGCTTCAACATC 2334  
780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAla 799  
2335 AACTTCAACATCACTAAGGACACAAAGCTTGTGCTGAATTGTGGCTCTGGAGAGTGAAGG 2394  
800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
2395 GGGTCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAAAAG 2454  
820 IleIleSerSerLeuAspProProCysArgArgGlyValaAspTrpArgThrLeuAlaGln 839  
2455 ATCATCCCTAGTCTGGACCCACCTCTGCAGCGGGGCGCCGAGCTGAGAGACTCTAGCCAG 2514  
840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859  
2515 AACTTCACTGGACAGCCATCTTAGCTTCTTTCCTCCAAAGCCAGCCCTTACAGCCATG 2574  
860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879  
2575 ATCTCAACCTATGGAGGCGACGCGACTTCCCCAACGCGCAACCTCGGCCAGCTGGCAGCA 2634  
880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
2635 GCTGTGGCCGAGCTGGGCGCAACCAAGATGCTGGCCCTCTTACCGTGTGGAGGCGGAGTGT 2694  
RESULT 5  
AAV52940  
ID AAV52940 standard; cDNA; 3014 BP.  
XX  
AC AAV52940;  
XX  
DT 25-MAR-2003 (revised)  
DT 21-DEC-1998 (first entry)  
XX  
DE Rat UNC-5 homologue unc5h-1 cDNA.  
XX  
KW UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;  
KW diagnosis; therapy; ds.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2697  
FT /\*tag= a  
XX  
PN WO9837085-A1.  
XX  
PD 27-AUG-1998.  
XX  
PF 19-FEB-1998; 98WO-US003143.  
XX

PR 19-FEB-1997; 97US-00808982.  
 XX (REGC ) UNIV CALIFORNIA.  
 PA Teasier-Lavigne M, Leonardo ED, Hinck L, Maseu M, Keinomasu K;  
 XX  
 XX  
 XX  
 XX  
 DR WPI; 1998-495364/42.  
 DR P-ESDB; AAW78898.  
 XX  
 PT Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and  
 PT the biopharmaceutical industry.  
 XX  
 PS Claim 7; Page 15-17; 32pp; English.  
 XX  
 CC This cDNA, termed unc5h-1, comprises a rat homologue of Caenorhabditis  
 CC elegans unc-5. Rat unc5h-1 and unc5h-2 (see AAV52942) cDNAs were isolated  
 CC from an E18 brain cDNA library. The predicted proteins (see AAW78898 and  
 CC AAW78900) show similarity with UNC-5. They are predicted to be involved  
 CC in cell migration and axon guidance, and are characterised as receptor  
 CC proteins for netrins. Gene expression is observed in regions where  
 CC differentiating neurons are undergoing axogenesis. Human unc5h-1 (see  
 CC AAV52941) and unc5h-2 (see AAV52943) cDNAs are also claimed. Vertebrate  
 CC UNC-5 proteins may be produced recombinantly from transfected host cells  
 CC by utilising these vertebrate UNC-5 nucleic acids. The invention also  
 CC provides unc-5 hybridisation probes and primers, vertebrate UNC-5-  
 CC specific binding agents such as specific antibodies, and methods of  
 CC making and using the subject compositions in diagnosis (e.g. genetic  
 CC hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g.  
 CC gene therapy to modulate vertebrate unc-5 gene expression) and in the  
 CC biopharmaceutical industry (e.g. as immunogens, reagents for modulating  
 CC cell guidance, reagents for screening chemical libraries for lead  
 CC pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI  
 CC field.)  
 XX  
 SQ Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,84e-304 Length: 3014  
 Score: 4545.50 Matches: 852  
 Percent Similarity: 96.78% Conservative: 19  
 Best Local Similarity: 94.67% Mismatches: 26  
 Query Match: 94.96% Indels: 3  
 DB: 2 Gaps: 3

US-09-970-944-2 (1-899) x AAV52940 (1-3014)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaLaTrpLeu 20  
 DB 1 ATGGCCGTCGCGCCCGCCCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCCGCTGGCTT 60

QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
 DB 61 CTGTGGTTCGGGTGCCAGCAGAGTGCACCGTGGCCAAATCCAGTGGCCGGTGCCAAACCCC 120

QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValIysAsnLysPro 60  
 DB 121 GACCTGCTGCCACATCTCTCTGTAGAGCTGTAGAGCGTGTACATTGTCAAGACACAGCCG 180

QY 61 ValLeuLeuValCysIysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
 DB 181 GTGTGTTGGTGGCAAGGCTGTGCTGCCACCCAGATCTTCTCAAGTGCATGGGAA 240

QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
 DB 241 TGGGTCCGCAGGTGCATACGTATTGAAACGACGACCCAGCACCGACGACGCGGATGGCCA 300

QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluIysValPheGlyLeuGlu 120  
 DB 301 ACCATGGAGTGGTATCAACGATATCGAGCAGCAGGTAGAGAAAGTGTGGGCTGGAG 360

QY 121 GluTyrTrpCysGlnCysValAlaLaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140  
 DB 361 GAATACTGTGGCAGTGTGTGGCATGGAGTCTCTCGGTACCAACCAAAAGTCAGAGGCC 420

QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
 DB 421 TACATCCGGATTGGCTATTTCGCAAGAACTTTGAGCAGGAGGCACTGCGCAAGAAAGTG 480

QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180  
 DB 481 TCATCTGGAGCAAGCATTTGTACTACCTTGTGCGCCCCCAGAGGAATCCCCCACTGAG 540

QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
 DB 541 GTGGAGTGGCTTCGAATGAGGACCTCGTGACCCCTCCCTCGATCCCAATGTGTACATC 600

QY 201 ThrArgGlnHisSerLeuValValArgGlnAlaAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
 DB 601 ACGGGGAGCAGACAGCTAGTGTGCGTCAGCGCCGCTGGCCGACACGGCCCACTACACC 660

QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
 DB 661 TGTGTGCCCAAGAACATCTGTAGCCCGTCGCCAAGCACCTCTGCAGCGGTCAITGTTTAT 720

QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
 DB 721 GTGAACGGTGGGTGGTGCAGCTGACCTGAGTGGTCCGTCCTGCAGCGCCAGCTGTGGCGT 780

QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
 DB 781 GGTGTGGCAGAAACGGAGCCGAGCTGCACCAACCCGCGACCTCTCAACGGGGGCGCCTTC 840

QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
 DB 841 TGTGAGGGGAGAAATGTC---CAGAAACAGCCTGCGCCACTCTGTGTCGCCAGTGGATGGG 897

QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
 DB 898 AGCTGGAGTTCTGTGGAGTAAGTGGTGCAGCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957

QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
 DB 958 CGCGAGTGTCTGTACCCAGCAGCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGCTGACCTG 1017

QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
 DB 1018 GACCCCGCAACTGTACAGTACCTGTGCTGCACACCGCTCTTGTGCCCGGAGGAGCTG 1077

QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
 DB 1078 GCTCTCTACATCGCCCTTGTGCTGTGGCTGTGTGCTCTTCTTGTGTGTGGCCCTT 1137

QY 381 IleLeuValTyrCysArgLysLysGluLeuAspSerValAlaAspSerSerIle 400  
 DB 1138 GGACTCATTTACTGTGCGCAAGAGAGGAGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197

QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
 DB 1198 CTACCTTCGGCTTTCAGCCCTGTGACATCAAGCCCGAGCAAGAGCAGCAACCCCACTG 1257

QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr 439  
 DB 1258 CTCACCATCCAGCCAGACCTCAGCACCACTACCACTACCACTACCACTACCACTACCACT 1317

QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
 DB 1318 AGGCAGGATGGAGCCAGCCCAAGTTCCAGCTCTCTAATGGTCACTGTGTGAGCCCACTG 1377

QY 460 GlyGlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGluGluPheVal 479  
 DB 1378 GGGAGTGGCGCCGACACGTTCAGCACCACTACCACTACCACTACCACTACCACTACCACT 1437

QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
 DB 1438 TCCCGCTCTCCACCCCAAAATACCTTTGTTCCCTCCCGCCGCGCACCAAGCAACATGGCC 1497

QY	500	TyrGlyThrPheAsnProGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519	Db	2575	ATCCTCAACCTATGGGAGGACGCGACTTCCCAACGCGCAACCTCGGCCAGCTGGCAGCA	2634		
Db	1498	TACGGGACCTTCAACTCTTCGGGGGGCGGCTGATGATCCCTAATACGGGGATCAGCCTC	1557	QY	880	AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899		
QY	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis	539	Db	2635	GCTGTGGCGGAGTGGGCCCAACAGATGCTGGCCTTTTCACGGTGTGGAGGCGCGAGTGT	2694		
Db	1558	CTCATACCCCGGATGCCATCCCGGAGGAAAGATCTACGAGATCTACCTCACACTGCAC	1617	RESULT 6					
QY	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559	AAK52261	AAK52261 standard; cDNA; 2907 BP.				
Db	1618	AAGCCAGAACGCTGAGGTGGCCCTAGCTGGCTGCAGACCCCTGCTGAGTCCAGTCGTT	1677	XX	AAK52261;				
QY	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579	XX	06-NOV-2001	(first entry)			
Db	1678	AGCTGTGGGCCCCAGAGTCTCTGCTCACCGGCGCAGTCATCTTGCATGGACCATGT	1737	DE	Human polynucleotide SEQ ID NO 806.				
QY	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer	599	XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;				
Db	1738	GGAGAGCCAGCCCTGACAGCTGGAGTCTGGCCTCAAAAGCAGTCTCTGCGAGGGCAGT	1797	KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;				
QY	600	TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln	619	KW	tissue growth factor; immunomodulatory; cancer; leukaemia;				
Db	1798	TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACTTCCCAACCTCTACTACTGCAG	1854	XX	nervous system disorder; arthritis; inflammation; ss.				
QY	620	LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly	639	OS	Homo sapiens.				
Db	1855	CTGGAGCGCGGGCTGCTATGTCTTTCAGGAGCAGCTGGCGCGCTTGTCCCTCGTAGGA	1914	XX	WO200157190-A2.				
QY	640	GluAlaLeuSerValAlaAlaAlaArgLeuLysLeuLeuPheAlaProValAla	659	XX	09-AUG-2001.				
Db	1915	GAGGCCCTCAGCTGGCTGCCACCAAGCGCTCAGGCTCTCTTGTTCCTCGGTGGCC	1974	XX	05-FEB-2001; 2001WO-US0004098.				
QY	660	CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu	679	XX	03-FEB-2000; 2000US-00496914.				
Db	1975	TGTACGTCCTTGAGTACAACTCCAGTGTACTGCCCTACAGCACCCACGAGCGCTC	2034	PR	27-APR-2000; 2000US-00560875.				
QY	680	LysGluValValGlnLeuGlyGlnLeuGlyGlnLeuIleGlnGluProArgVal	699	PR	20-JUN-2000; 2000US-00598075.				
Db	2035	AAGGAGTGGTGCAGCTGGAAGCAGCTAGTGGAGCAGCTGATCCAGGAGCCTCGGCTC	2094	PR	19-JUL-2000; 2000US-00620325.				
QY	700	LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer	719	PR	01-SEP-2000; 2000US-00654936.				
Db	2095	CTGCACCTCAAGACAGTTACCACAACTAGCTCTCTCCATCCAGCAGCTGCCAGCTCC	2154	PR	15-SEP-2000; 2000US-00663561.				
QY	720	LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn	739	PR	20-OCT-2000; 2000US-00693325.				
Db	2155	CTGTGGAAGCAAGCTACTTGTACGTACACGAGAGATCCCTTTTACCATCTGGAAC	2214	PR	30-NOV-2000; 2000US-00728422.				
QY	740	GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer	759	XX	(HYSE-) HYSEQ INC.				
Db	2215	GGCACCAGCAGTATCTGACCTGCACCTTCCCTGGAGCGCATCAACGCCAGCACCCAGC	2274	XX	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;				
QY	760	AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle	779	XX	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;				
Db	2275	GACCTGGCTGCAAGGTGGTGGGAGTGGGAGGAGATGGCGAGCTTCAACATC	2334	XX	Xue AJ, Yang Y, Wejhrman T, Goodrich R;				
QY	780	AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla	799	XX	WPI; 2001-476283/51.				
Db	2335	AACITCAACATCCTCAAGACACAAAGTTTGTGAAATTTGGCTCTCGAGAGTGAAGGG	2394	XX	P-PSDB; AAM79128.				
QY	800	GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys	819	XX	Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.				
Db	2395	GGGTCCCGACCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAAG	2454	XX	Claim 1; Page 2691-2694; 6221pp; English.				
QY	820	IleLeuSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln	839	XX	The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication				
Db	2455	ATCATCGCAGCTGGACCCACCTGCAGCGGGCGCGAGCTGAGAACTCTAGCCAG	2514	XX	Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;				
QY	840	LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet	859	XX	Alignment Scores:				
Db	2515	AAACTTCACCTGGACAGCATCTTATGCTTCTTGTGCTTCAAGCCCGAGCCCTACAGCCATG	2574	XX	Pred. No.: 1.84e-296 Length: 2907				
QY	860	IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	879	XX	Score: 4434.00 Matches: 853				
				XX	Percent Similarity: 90.78% Conservative: 4				

Best Local Similarity: 90.36%		Mismatches: 17
Query Match: 92.63%		Indels: 70
DB: 4		Gaps: 7
US-09-970-944-2 (1-899) x AAKS2261 (1-2907)		
QY	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
DB	148	ATGACGAGGCGTCCCTCCCTG-----ATGGCGGCGAGACAG 183
QY	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
DB	184	CACGATGGAGCGCCGACAGAGTGCACCGTGGCCAAACAGTGCCTGGTGCACACCG 243
QY	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
DB	244	GACCTGCTTCCCACTTCCCTGGTGGAGCCGAGATGTACATCGTCAAGAACAAAGCCA 303
QY	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
DB	304	GTGCTGCTGTGTGCAAGGCGCGTCCCGCCACGACAGATCTTTCAGTGCACAGGGGAG 363
QY	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
DB	364	TGGGTGCGCAGGTGGACACAGTATCGAGCGAGCAGACAGACGGAGCAGTGGGCTGCC 423
QY	101	ThrMetGluValArgIleAsnValSerArgGlnValGlnLysValPheGlyLeuGlu 120
DB	424	ACCATGAGGTCCGATTAATGCTCAAGGAGCAGGTTCGAGAGGTTCGGGCTGGAG 483
QY	121	GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
DB	484	GAATACTGTGTGCAGTGGTGGATGGAGCTCCTCGGCGACACCAAGAGTCAAGAGGCC 543
QY	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGlnGluProLeuAlaLysGluVal 160
DB	544	TACATCGCATAGCCTATTTCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGAGGTG 603
QY	161	SerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGlu 180
DB	604	TCCTTGAGAGGAGGAGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY	181	ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
DB	664	GTGAGTGGCTCGGAACAGAGGACCTGGTGGACCCCGTCCCTGGACCCCAATGATATC 723
QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
DB	724	ACGCGGAGCACAGCTGGTGGTGGACAGGCGCGCTTCTGACACGGCCCAACTACACC 783
QY	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr 240
DB	784	TGCGTGGCCNAGAACATCGTGGCAGCTGCGCGGAGCGCTCCCTGCTGTCTGCTCTAC 843
QY	240	-----240
DB	844	GGTGGGCGCGGACTCCCTGCTCACAGGAGAGGACATGCGGTGCGCCCTGGCGAGTGAC 903
QY	241	-----ValAsnGlyGlyTrpSerThrTrpThrGluTrp 251
DB	904	ATGTGCTGTCTCTGTGCGGCCAGTGAACGGTGGGTGGTGGAGCCGAGTGG 963
QY	252	SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271
DB	964	TCCGTCTGACGCCAGCTGTGGGCGCGCTGGCAGAAACAGGAGCGGAGCTGCACCAAC 1023
QY	272	ProAlaProLeuAsnGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291
DB	1024	CCGGCGCTCTCAACGGGCGCTTCTGTGAGGGCGAGATGTC---CAGAAAAAGGCC 1080
QY	292	SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311
DB	1081	TGCGCCACCGCTGTGCCAGTGGAGCGGAGCTGGAGCGCGTGGAGCAAGTGGTGGCGCTGT 1140

QY	312	GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331
DB	1141	GGGCTGGACTGCACCCACTGCGGAGCCGTGAGTGCTCTGACCCAGACCCCGAACGGA 1200
QY	332	GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351
DB	1201	GGGAGAGTGGCAGGCGACTGACCTGGACACCCGCAACTGTACCATGACCTCTGTGTA 1260
QY	352	His-----SerAlaSerGly 356
DB	1261	CACAACCTCTACACCCCTGCCCCCACCAGGCGCATGCTGCTCCCGCAGCTGCTTCTGGC 1320
QY	357	ProGluAspValAlaLeuTyrValGlyIleAlaValAlaValCysLeuValLeuLeu 376
DB	1321	CCTGAGACGTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTGCTGCTGCTGCTGCTG 1380
QY	377	LeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAla 396
DB	1381	CTGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
QY	397	AspSerSerIleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAsp 416
DB	1441	GACTCGTCCATTCTCACCTCAGCTTCAGCCCGTCCAGCATCAAGCCAGCAAGAGCAGAC 1500
QY	417	AsnProHisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGly 435
DB	1501	AAACCCCATCTGCTCCACCATCCAGCCGCGACTCAGCACCAACCCACCACTACCAAGGC 1560
QY	436	SerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeu 455
DB	1561	AGTCTGTGCTCCCGGAGGATGGGCGCCAGCCCAAGTTCAGCTCACCATTGGGACCTG 1620
QY	456	LeuSerProLeuGlyGlyArgHisThrLeuHisHisSerSerSerProThrSerGluAla 475
DB	1621	CTCAGCCCGCTGGTGGCGCGCCACACACTGCACACACAGCTCTCCCTCTGAGGCC 1680
QY	476	GluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThr 495
DB	1681	GAGAGTGTCTCTCCCGCTCTCCAGCCAGAACTACTTCCGCTCCCTGCGCGGAGGCACC 1740
QY	496	SerAsnMetThrTyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThr 515
DB	1741	AGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACA 1800
QY	516	GlyIleSerLeuLeuLeuProAspAlaIleProArgGlyLysIleTyrGluIleTyr 535
DB	1801	GGAAATCAGCTCTCTCTATCCCGCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTAC 1860
QY	536	LeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeu 555
DB	1861	CTCAGCTGCAAGCGGAGAGAC-----1884
QY	556	SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla 575
DB	1885	-----GTGAGCTGTGAGACCCCTGCGCTCTGCTCACCGGCGAGTATCTCTGGCT 1935
QY	576	MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSer 595
DB	1936	ATGAGCACTGTGGGAGCCAGCCCTGACAGCTGAGCTGCGCTCAAAAAGAGAGTGC 1995
QY	596	CysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu 615
DB	1996	TGCGAGGGAGCTGGAG---GATGTGCTGCACCTGGGCGAGAGCGCGCTCCCACTC 2052
QY	616	TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe 635
DB	2053	TACTACTGCAGCTGGAGCGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2112
QY	636	AlaLeuValGlyGluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPhe 655
DB	2113	GCCCTGGTGGAGAGGCCCTCAGCGTGGCTGCGCGCAAGCGCTCAAGCTGCTTCTGTTT 2172



```
QY 656 AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr 675
DB 2173 GCGCCGGTGGCCCTGCACCTCCCTCGAGTACAATCCGGGTCTACTCCCTGCATGACACC 2232
QY 676 HisAspAlaLeuLysGluValValGlnLeuGlnLysGlnLeuGlyGlyGlnLeuIleGln 695
DB 2233 CACGATGCATCAAGGAGGTGGTGCAGTCGAGAGACAGCTGGGGGACACGCTGATCCAG 2292
QY 696 GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp 715
DB 2293 GAGCCAGGGTCTCTGCACCTTCAAGGACGATGACACAACTGGCGCTATCCATCCACGAT 2352
QY 716 ValProSerSerLeuTyrLysSerLysLeuValSerTyrGlnGluIleProPheTyr 735
DB 2353 GTGCCAGCTCCCTGTGGAGAGTAAGCTCTTGTACGTACACAGGAGATCCCTTTTAT 2412
QY 736 HisIleTyrAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer 755
DB 2413 CACATCTGGAATGGCAGCAGCGGTACTTGACCTGCACCTTACCTTGGAGCGTGTCCAGC 2472
QY 756 ProSerThrSerAspLeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGln 775
DB 2473 CCCAGCAGTAGTACCTGGCTGCAGCTGTGGTGTGGCAGGTGGAGCGGCGGCGAG 2532
QY 776 SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeu 795
DB 2533 AGCTTACGATCAACTTCAACATCACCAGAGACCAAGGTTTGTGAGCTGTGGCTCTG 2592
QY 796 GluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeu 815
DB 2593 GAGAGTGAAGCGGGTCCAGCGCTGTGGGCCAGTGCCTTCAAGATCCCTTCTC 2652
QY 816 IleArgGlnLysIleLeuSerSerLeuAspProCysArgArgGlyAlaAspTyrArg 835
DB 2653 ATTGGCAGAGATAATTTCCAGCTGCAGCCACCTGTAGCGGGGTGGCGACTGGCGG 2712
QY 836 ThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSer 855
DB 2713 ACTGTGCCCGCAGAACTCCACCTGGACAGCATCTCAGTCTTGTGCTCCAGCCGAGC 2772
QY 856 ProThrAlaMetIleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSer 875
DB 2773 CCCACAGCCATGATCTCTCAACCTGTGGAGCGCGGCACCTTCCCAACGGCAACCTCAGC 2832
QY 876 GlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSer 895
DB 2833 CAGCTGGCTCAGCAGTGGTGGAGTGGCTGGCGGAGCCAGCAGCTGGCTTTCACAGTGTGC 2892
QY 896 GluAlaGluCys 899
DB 2893 GAGGCTGAGTGC 2904
RESULT 7
ABK15169
ID ABK15169 standard; DNA; 3580 BP.
XX
AC ABK15169;
XX
DT 23-APR-2002 (first entry)
XX
DE Human REPTR 1 cDNA sequence.
XX
KW REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;
KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;
KW antiallergic; antibody; immunogen; endometriosis;
KW gastrointestinal disorder; gastritis; oesophageal carcinoma;
KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;
KW endocrine disease; hypothalamus disorder; Kallman's disease;
KW autoimmune disease; inflammatory disease; infertility; receptor;
KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
KW osteoarthritis; diabetes mellitus; multiple sclerosis;
KW systemic lupus erythematosus; cell proliferative disorder; cancer;
KW developmental disorder; Duchenne muscular dystrophy; gene;
```

```
KW Becker muscular dystrophy; neurological disorder; epilepsy;
KW Alzheimer's disease; Huntington's disease; reproductive disorder; ss.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 4..2532
XX /tag= a
XX /product= "REPTR1 protein"
XX
XX WO200198354-A2.
XX
XX 27-DEC-2001.
XX
XX 21-JUN-2001; 2001WO-US019942.
XX
XX 21-JUN-2000; 2000US-0214027P.
XX 25-AUG-2000; 2000US-0228045P.
XX 12-DEC-2000; 2000US-0255104P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;
XX Lal P, Policky JL, Azimzai Y, Lu DAM, Graul R, Yao MG, Burford N;
XX Hafalia AJA, Baughn MR, Bandman O, Patterson C, Yang J, Xu Y;
XX Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;
XX
XX WPI; 2002-090432/12.
XX P-PSDB; AAU17818.
XX
XX Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in
XX the diagnosis, treatment and prevention of gastrointestinal (e.g.
XX gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell
XX proliferative (e.g. cancer) disorders.
XX
XX Claim 57; Page 142-143; 157pp; English.
XX
XX This invention relates to twelve human receptors cDNA sequences referred
XX to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The
XX proteins of the invention may have antiinflammatory, cytostatic, active
XX immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular, active
XX general, anticonvulsant, nootropic, neuroprotective, antiallergic
XX activities. The sequences of the invention may be used to produce REPTR
XX agonists or antagonists, and the protein sequences may be used to raise
XX anti-REPTR antibodies. These molecules and the REPTR polynucleotides and
XX polypeptides of the invention are useful in the diagnosis, treatment and
XX prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,
XX Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine
XX (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory
XX (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,
XX allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,
XX systemic lupus erythematosus), cell proliferative (e.g. cancer),
XX developmental (e.g. Duchenne and Becker muscular dystrophy), neurological
XX (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and
XX reproductive (e.g. infertility, endometriosis) disorders. Numerous other
XX examples of each disorder are given in the specification. The present
XX sequence represents the human REPTR1 cDNA sequence of the invention
XX
XX Sequence 3580 BP; 670 A; 1233 C; 1025 G; 652 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1.41e-292 Length: 3580
Score: 4379.50 Matches: 838
Percent Similarity: 93.22% Conservative: 1
Best Local Similarity: 93.11% Mismatches: 2
Query Match: 91.49% Indels: 59
DB: 6 Gaps: 3
```

US-09-970-944-2 (1-899) x ABK15169 (1-3580)

```
Qy 1 MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAlaTyrLeu 20
DB 4 ATGGCCGCTCCGGCCCGCCCTGTGGCCAGCGCTCTCTGGGCATAGTCTCGCGCTTGGCTC 63
```

QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
DB 64 CGCGGCTCGGGTGCCACGAGTGCACCGTGGCCAAACCAGTGGCTGGTGCCAAACCG 123  
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60  
DB 124 GACCTGCTTCCCCACCTTCCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACAGCCA 183  
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
DB 184 GTCTGCTGTGTGCAAGGCCGTGCCCGCCACGACAGATCTTCTCAAGTGCACAGGGGAG 243  
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
DB 244 TGGGTGGCCAGTGGACACGTAATGCTCAAGGCAGCAGGTTCGAGAAAGGTGTTTCGGGCTGGAG 303  
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
DB 304 ACCATGGAGTCCGCATTAATGCTCAAGGCAGCAGGTTCGAGAAAGGTGTTTCGGGCTGGAG 363  
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140  
DB 364 GAATATCGGTGCGAGTGGCATGGAGCTCCTCGGGCACCAAGAGTCAAGAGGCC 423  
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
DB 424 TACATCCGATAGCTATTGTCGCAAGAACTTCGACGAGAGCGCTGGCCAAAGAGGGT 483  
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProAlaGlu 180  
DB 484 TCCTCGAGCAGGCGCATCGTGTGCTGCCGTCCAGCGAGGCGATCCCTCCAGCCGAG 543  
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerSerLeuAspProAsnValTyrIle 200  
DB 544 GTGGAGTGGCTCGGAACAGAGACCTGGTGGACCCCGTCCCTGGACCCCAATGTATACATC 603  
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
DB 604 ACSCGGAGCACAGCTGTGTGTCGACAGCCGCGCTTCTGACAGCGCCAACTACACC 663  
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
DB 664 TGGCTGGCCAAAGAACATCGTGGCACCTGCGCGAGCGCCTCCGCTGCTCATCGTCTAC 723  
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
DB 723 ----- 723  
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
DB 723 ----- 723  
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
DB 724 -----GTGGACGGC 732  
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
DB 733 AGCTGAGCCCGTGGAGCAAGTGGTGGCTGTGGGCTGGACTGGACCCCACTGGCGGAGC 792  
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340  
DB 793 CGTGAAGTCTTACCCAGCACCCCGAACCGAGGGGAGGAGTGGCCAGGGCATGACCTG 852  
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
DB 853 GACACCCGCAACTGTACCACTGACCTGTGTATACACACTGCTTCTGGCCCTGAGGACGTG 912  
QY 361 AlaLeuTyrValGlyIleuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
DB 913 GCCCTATGTGGGCGCTCATCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCT 972

QY 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400  
DB 973 ATCTCTGTTTATTGCGGAGAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1032  
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
DB 1033 CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAAGCAAGACCAACCCCATCTG 1092  
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439  
DB 1093 CTCACCATCCAGCCGACCTCAGCACCAACCAACCACTTACCAAGGCGAGTCTCTGTCCC 1152  
QY 440 ArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
DB 1153 CGGACAGATGGGCCAGCCCCAAGTTCAGCTCACAATGGGCACCTGTCTACGCCCTG 1212  
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479  
DB 1213 GGTGGCGGCGCCACACACTGCACCACTCTCCACCTCTGAGGCGGAGGAGTTCGTC 1272  
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
DB 1273 TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGCACCAACATGACC 1332  
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
DB 1333 TATGGACCTTCACTTCTCGGGGCGCGCTGATGATCCCTAATACAGGAATCAGCCTC 1392  
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539  
DB 1393 CTCATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTACCTACGCTGCAC 1452  
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
DB 1453 AAGCCGGAAGACGTAGGTTCGCTTGTGGCTGTGCAGACCTGTGATGCCATCGTT 1512  
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
DB 1513 AGCTGTGACCCCTGGCGTCTGTCTACCGGCCAGTCTCTGTGCTATGGACCACTGT 1572  
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599  
DB 1573 GGGGAGCCCGACCTCAGACCTGGAGCTGGGCTCAAAAAGCAGTCTGCGAGGGCAGC 1632  
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrCysGln 619  
DB 1633 TGGGAG---GATGTGTGCACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCCAG 1689  
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluLeuGlyArgPheAlaLeuValGly 639  
DB 1690 CTGGAGCCAGTGCCTGCTTTCACCGAGCAGCTGGCGCGCTTGGCCCTGGGTGGGA 1749  
QY 640 GluAlaLeuSerValAlaAlaLysArgLeuLeuLeuLeuPheAlaProValAla 659  
DB 1750 GAGGCCCTCAGCGTGGCTGCCCGCAAGCGCTCAAGCTGCTTCTGTTCGCGCGGTGGCC 1809  
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
DB 1810 TGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCTGCAATGACACCAACCATGCACTC 1869  
QY 680 LysGluValValGlnLeuGlyGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699  
DB 1870 AAGGAGTGGTGCAGCTGGAGAGCAGCTGGGGGAGACAGCTGATCCAGAGGCCACGGTTC 1929  
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
DB 1930 CTGCATCTCAAGACAGTATTACCAACCTGGCGCTATCCATCCACGATGTGCCAGCTCC 1989  
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
DB 1990 CTGTGAAGAGTAAGCTCCTTGTGCTACGTACCAAGAGATCCCTTTTATCACATCTGAAT 2049  
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759

```

Db      2050 GGCACGAGCGGACTTGGCACTTCACCTTCACCTGGAGCGGTGCACCCCGACAGCTAGT 2109
QY      760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db      2110 GACCTGGCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGGACGGCAGAGCTTCAGCATC 2169
QY      780 AspPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db      2170 AACTTCAACATCAACAGGACACAAAGGTTTGCTGAGCTGTGGCTCTGGAGAGTGAAGCG 2229
QY      800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db      2230 GGGGTCCACGCTGTGGGCCCGCCAGTGCCTTCAAGATCCCTTCCTCAITCGGCAGAG 2289
QY      820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db      2290 ATAATTTCCAGCTGGACCCACCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCAG 2349
QY      840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db      2350 AAATCTCACTGGACGCCATCTCAGCTTCTTTGCCCTCCAAAGCCCGACGCCACAGCCATG 2409
QY      860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db      2410 ATCTTCACCTGTGGAGGGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2469
QY      880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db      2470 GCAGTGGCTGGACTGGCCAGCCAGACGCTGGCTCTTCACAGTGTGGAGGCTGAGTGC 2529

RESULT 8
AAV52941
ID      AAV52941 standard; cDNA; 1787 BP.
XX      AC
XX      AAV52941;
XX      25-MAR-2003 (revised)
XX      21-DEC-1998 (first entry)
XX      Human UNC-5 homologue unc5h-1 cDNA.
XX      DE
XX      UNK-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;
XX      diagnosis; therapy; ds.
XX      OS
XX      Homo sapiens.
XX      FH
XX      CDS
XX      3..1670
XX      Location/Qualifiers
XX      /*tag= a
XX      /transl_except= (pos:21..22, aa:Xaa)
XX      /transl_except= (pos:200..203, aa:Ile)
XX      /transl_except= (pos:771..772, aa:Ser)
XX      /transl_except= (pos:785..786, aa:Leu)
XX      /transl_except= (pos:1078..1079, aa:Xaa)
XX      /transl_except= (pos:1098..1099, aa:Xaa)
XX      /transl_except= (pos:1106..1107, aa:Xaa)
XX      /transl_except= (pos:1621..1622, aa:Gly)
XX      /note= "these codons have apparent 1 nucleotide deletions
XX      or insertions, which alter the reading frame"
XX      PN
XX      WO9837085-A1.
XX      PD
XX      27-AUG-1998.
XX      PF
XX      19-FEB-1998; 98WO-US0031143.
XX      PR
XX      19-FEB-1997; 97US-00808982.
XX      PA
XX      (REGC ) UNIV CALIFORNIA.
XX      PI
XX      Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;
XX

```

```

DR      WPI; 1998-495364/42.
DR      P-PSDB; AAW78899.
XX
PT      Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
PT      the biopharmaceutical industry.
XX
PS      Claim 7; Page 17; 32pp; English.
XX
CC      This partial cDNA, termed unc5h-1, comprises a human homologue of
CC      Caenorhabditis elegans unc-5, unc5h-1 and unc5h-2 (see AAV52943) cDNAs
CC      were isolated from an embryonic brain cDNA library. The predicted
CC      proteins (see AAW78899 and AAW78901) show similarity with UNC-5. They are
CC      predicted to be involved in cell migration and axon guidance, and are
CC      characterised as receptor proteins for netrins. Gene expression is
CC      observed in regions where differentiating neurons are undergoing
CC      axogenesis. Rat unc5h-1 (see AAV52940) and unc5h-2 (see AAV52942) cDNAs
CC      are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
CC      from transfected host cells by utilising these vertebrate UNC-5 nucleic
CC      acids. The invention also provides unc-5 hybridisation probes and
CC      primers, vertebrate UNC-5-specific binding agents such as specific
CC      antibodies, and methods of making and using the subject compositions in
CC      diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5
CC      transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5
CC      gene expression) and in the biopharmaceutical industry (e.g. as
CC      immunogens, reagents for modulating cell guidance, reagents for screening
CC      chemical libraries for lead pharmacological agents, etc.). (Updated on 25
CC      -MAR-2003 to correct PI field.)
XX
XX      SQ      Sequence 1787 BP; 349 A; 603 C; 491 G; 344 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      158-176      Length:      1787
Score:          2694.00      Matches:    543
Percent Similarity: 97.14%      Conservative: 1
Best Local Similarity: 96.96%      Mismatches: 11
Query Match:     56.28%      Indels:     11
DB:              2          Gaps:        2

US-09-970-944-2 (1-899) x AAV52941 (1-1787)
QY      344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db      3 AACTGTACAGTGCACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61
QY      364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuValLeuVal 383
Db      62 GTGGGCTCATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY      384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
Db      122 TATTGCCGGAAGAGGAGGGCTGGACTCAGATGTGGCTGACTGCTCCATCTCACCTCA 181
QY      404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIle 423
Db      182 GGCTTCCAGCCCTCAGCATCTTAAGCCAGCAAAAGCAGACAAACCCCATCTGCTCCACAT 241
QY      423 eGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThrThrThrThr 442
Db      242 CCAGCCGAGCTCAGACACACACACACACACACACACACACACACACACACACACACACAGGA 301
QY      442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGly 462
Db      302 TGGGCCCGAGCCCAAGTTCAGCTCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 361
QY      462 YArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
Db      362 CCGCCACACACTGCACACAGCTCTCCACCTCTGAGGGCCGAGAGTCTGCTCTCCGCCCT 421
QY      482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrThrGlyTh 502
Db      422 CTCACCCAGAACTACTTCGCTCCCTCCGCCGAGGACCCAGCAACATGACCTATGGGAC 481
QY      502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522

```

Db 482 CTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGATCAGCTCTCTCATCCC 541  
 QY 522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProG1 542  
 Db 542 CCAGATGTCATACCCGAGGAGATCTATGAGATCTACCTCAGCTGCACAGCCGGA 601  
 QY 542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysG1 562  
 Db 602 AGAGTGTAGGTTGCCCTAGCTGGCTGTGACACCCCTGCTGAGTCCCATCTTACGTGTGG 661  
 QY 562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582  
 Db 662 ACCCCTGTGGCGTCTCTCACCAGGAGATCTATGAGTATGACCTATGACCTGTGGAGCC 721  
 QY 582 oSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluG1 602  
 Db 722 CAGCCTGTAGAGTGGAGCCCTGAGCCCTCAAAAAGCAGTCTGCGAGGAGC-TGGGAG-- 778  
 QY 602 nAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622  
 Db 779 -GATGT-CTGCACCTGGCGAGAGGCGCCCTCCACCTCTACTTGCCTGAGTGGAGGC 836  
 QY 622 aSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642  
 Db 837 CAGTGTCTGTCTTACCAGCAGCTGGCGCGCTTTTGCCTGTGTGGAGAGCCCT 896  
 QY 642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSe 662  
 Db 897 CAGCGTGGCTGGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGGCTGGCCCTGCACCTC 956  
 QY 662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682  
 Db 957 CTTGAGTCAACATCCGGGTCTACTGCTGTGATGACCCAGATGCATCTCAAGAGGT 1016  
 QY 682 lValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPh 702  
 Db 1017 GGTGACGTGGAGAGCAGCTGGGGGACAGCTGATCCAGAGCCACGGGTCTGCACCT 1076  
 QY 702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722  
 Db 1077 -AAGGACAGTTACCAACACCTGCG-CTATC-ATCCACGATGTGCCAGCTCCCTGTGGAA 1133  
 QY 722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742  
 Db 1134 GATAGACTCTTGTAGCTACAGAGATCCCTTTTATCATCTGGAAATGGACGCA 1193  
 QY 742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762  
 Db 1194 GCGGTACTTGCACTGCACCTTCACTCCCTGGAGCGTGTGAGCCAGCACTAGTCACTGGC 1253  
 QY 762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782  
 Db 1254 CTGCAAGCTGTGGGTGTGGCAGTGGAGGCGGAGCGGAGAGCTTCAATCAACTTCAA 1313  
 QY 782 nIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAlaGlyValPr 802  
 Db 1314 CATCACAGGACACAGGTTTGTGAGCTGTGCTGTGAGAGTGTGAGAGTGTGAGGTTCC 1373  
 QY 802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleLeSe 822  
 Db 1374 AGCCCTGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCTTCCATTCGCGAGAGATATTC 1433  
 QY 822 rSerLeuAspProProCysArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHi 842  
 Db 1434 CAGCCTGGACCCACCTCTAGGCGGGGTGCGGACTGTGCGGACTGTGCGCCAGAACTCCA 1493  
 QY 842 sLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAs 862  
 Db 1494 CTTGGACAGCCATCTCAGCTTCTTTGCTCCACAGCCCGAGCCCATGATCTCTCAA 1553  
 QY 862 nLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAl 882

Db 1554 CTTGTGGAGGCGGCGCACTTCCCCAACGCAACCTCAGCCAGCTGCTGCAGCAGTGGC 1613  
 QY 882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899  
 Db 1614 TGGGACTGGCCAGCAGGACGGTGGCTTCTTTTACAGTGTTCGGAGGCTGAGTGC 1667  
 RESULT 9  
 AAS75738  
 ID AAS75738 standard; cDNA; 3646 BP.  
 XX  
 AC AAS75738;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #11542.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR P-PSDB; ABG11551.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1; SEQ ID NO 11542; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have application of mutations  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3646 BP; 930 A; 917 C; 921 G; 877 T; 0 U; 1 Other;

Alignment Scores:  
 Pred. No.: 1e-174 Length: 3646  
 Score: 2673.50 Matches: 507  
 Percent Similarity: 72.40% Conservative: 154  
 Best Local Similarity: 55.53% Mismatches: 221

Query Match:	55.85%	Indels:	31
DB:	5	Gaps:	12
US-09-970-944-2 (1-899) x AAS75738 (1-3646)			
QY	9	ProAlaLeuLeuGlyIleValLeuAlaAaTrpLeuArgGlySerGlyAlaGlnGlnSer	28
DB	230	CTTGGCCCTG-----GCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCGCCCAAGATGAT	283
QY	29	-----AlaThrValAlaAaenProValProGlyAlaAaenProLeuLeuLeuProHis	45
DB	284	GACTTTTTTCATGAACCTCCCAAGAACTTTTCCTTCCTGATCCACTGAGCCCTCTGCCACAT	343
QY	46	PheLeuValGluProGluAspValTyrIleValIysAsnIysProValLeuValCys	65
DB	344	TTCCCTATTGAGCGCTGAAGAGCTTATATTGTGAAGAAATAGCTGTGAACCTGTACTGT	403
QY	66	LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal	85
DB	404	AAAGCAAGCCCTGCCACCCAGATCTATTTCAGTGTAAATAGTGAATGGGTTTCATCAGAAG	463
QY	86	AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg	105
DB	464	GACCACATAGTAGATAAAGAGTAGATGAACCTTCGGTCTCATTTGCCGGGAAGTGAGC	523
QY	106	IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln	125
DB	524	ATTGAGATTTCGCGCCAGCAAGTGAAGAACTCTTTGGACCTGAAGAAITACTGTGGCCAG	583
QY	126	CysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla	145
DB	584	TGTTGGCCCTGGAGCTCCGCGGTACCACAAAGACCGGAAGCGGTATGTGCGCAATTGCA	643
QY	146	ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly	165
DB	644	TATCTACGGAAGACATTTCGAGGAGAACCCCTAGGAAGAAAGTGTCTTTGGAACAGAA	703
QY	166	IleValLeuProCysArgProGluGlyIleProProAlaGluValGluTrpLeuArg	185
DB	704	GTCTTACTCAGTGTGCACCACTGAAGGATCCAGTGGCTGAGGTGGAATGTTGAAA	763
QY	186	AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer	205
DB	764	AATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTATATTACTATTGTATCAACAC	823
QY	206	LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn	225
DB	824	CTCATCAAAAGCAGGCGCCGACCTCTCTGATACCTGCAAAATTACACCTGTGTGCCAAAC	883
QY	226	IleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrp	245
DB	884	ATTGTTGCCAAGAGAAAAGTACAACGCGCACTGTCTATGTCTATGTCAACGGTGGCTGG	943
QY	246	SerThrTrpThrTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg	265
DB	944	TCCACCTGGACGAGTGTCTGTGTGTAAACAGCGCTGTGGACGAGGTATCAGAAACGT	1003
QY	266	SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlnAsn	285
DB	1004	ACAAGACTTGTATCAACACCGGCACCACTCAATGGGGGTGCCCTTCTGTGAAGGCGCAGT	1063
QY	286	ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrp	305
DB	1064	GTG---CAGNAATAGCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGACGCCATGG	1120
QY	306	SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp	325
DB	1121	AGCAAGTGTGTCTATTGTGAACTGAGTGACACCACTGGCGCAGGAGGAGTGACGCGC	1180
QY	326	ProAlaProArgAsnGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys	345
DB	1181	CCAGCCCCCAAGATGGGCAAGGACTGCGACGGCCCTGCTCTTGTCAATCCAGAACTGC	1240
QY	346	ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly	365
DB	1241	ACTGATGGGCTTTGCAATGCAGACTGCTCCTGATTCAGATGATGCTCTCTATGTTGGG	1300
QY	366	Leu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeuValLeuValTyr	384
DB	1301	ATTGTGATGACAGTGTGCTGCTGGCGATCTCTGTAGTTGTGGCTGTTGTTGTAT	1360
QY	385	CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGly	404
DB	1361	CGAAAGAATCATCGTCACTTTGATTCAGATATTATTGACTCTTCGGCAGCTCAATGGGGC	1420
QY	405	PheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThrIleGln	424
DB	1421	TTTCARCTGTGAACATCAAGGAGCAAGACAGAT-----CTGCTGGCTGTATCCC	1471
QY	425	ProAspLeu---SerThrThrThrThrThrGlnGlySerLeuCysProArgGlnAspGly	443
DB	1472	CCAGACCTCAGTCAGCTGCAGGCCATGTACAGAGGACCTGTCTATGCCCTGCATGAC---	1528
QY	444	ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyArg	463
DB	1529	GTCTCAGACAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCTGAAA	1588
QY	464	HisThrLeuHisHisSer-----SerProThrSerGluAlaGluGluPheVal	479
DB	1589	ATCAAGTGTACAAACCTCAGCTGCTGCTCCCCCAAGATGACCTCTCTGATTACG	1648
QY	480	SerArgLeuSer-----ThrGlnAsnTyrPhe-----	488
DB	1649	TCCAAGCTGCTCCCTCAGATGACCCAGTCGTTCTGGAGAATGAAGCCCTCAGCCTGAAG	1708
QY	489	---ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPheAsn	505
DB	1709	AACGAGCTGTACAAAGGAGACATGATCCATCTGTACCCGATTTGGCAGCTTCAACTCG	1768
QY	506	LeuGlyClyArgLeuMetIleProAsnThrGlyIleSerLeuLeuLeuProAspAla	525
DB	1769	CTGGAGGTCACTTATGTTCCCAATTCAAGAGTCAGCTTGTGATTCCTCGCTGGGGCC	1828
QY	526	IleProArgLysLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArg	545
DB	1829	ATTCCTCCCAAGGAGAGTCTACGAAATGTATGTACTGTACACAGGAAAGAACTATCAGG	1888
QY	546	LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly	565
DB	1889	CCACCATGGATGACTCTCAGACACTTTTGACCCCTGTGTGTGAGCTGTGGGCCCGCCAG	1948
QY	566	ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp	585
DB	1949	GCTCTGCTACCCGCGCGCTGCTCTCTCACTATGCATCCTGCGCAGACCCCAATACCGAG	2008
QY	586	SerTrpSerLeuArgLysLysGlnSerCysGlySerTrpGluGlnAspValLeu	605
DB	2009	GACTGAAATAATCTGCTCAAGAACCCAGGACAGCAGCAGTGGAG---GATGTGGTG	2065
QY	606	HisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCys	625
DB	2066	GTGGTGGGGAGAAAACCTTCACACCCCTGCTGTATTAAGCTGGATGAGAGGGCTGC	2125
QY	626	TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAla	645
DB	2126	CACATCTCACAGAGAACCTCAGCACCTAGCCCTGGTAGGACATTCACCCACCAAGCG	2185
QY	646	AlaAlaLysArgLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr	665
DB	2186	GCTGCAAAAGCGCTCAAGCTGGCCATCTTTGGCGCCCTGTGCTGCTCTCGCTGGAGTAC	2245
QY	666	AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu	685
DB	2246	AGCATCCGAGTCTACTGTCTGGATGACCCAGGATGCCCTGAGGAAATTTTACATCTT	2305
QY	686	GluLysGlnLeuGlyGlnLeuLeuGlnProArgValLeuHisPheLysAspSer	705



QY 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76  
DB 244 AAGAACAAGCTGTGGAGCTCCGTCGCGCGCTTCCCGCCACACAGATCTACTTCAAG 303  
QY 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96  
DB 304 TGCACGGCGAGTGGGTACGCCAGAACGACCACTCACAGGAGGCGCTGGATGAGGCC 363  
QY 97 Ser-----GlyGluProThrMetGluValArgIleAsnValSerArgGln 111  
DB 364 ACCCTGGGGCGCGGGCGCGCTCGCGGTGCGGAGGTGAGATCGAGATCGAGGTGTCGCGCAG 423  
QY 112 GluValGlnLysValPheGluLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSer 131  
DB 424 CAGGTGAGGAGCTCTTTGGGTGGAGATTACTGGTGGCAGTGGTGGCTGGAGCTCC 483  
QY 132 SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe 151  
DB 484 CGGGCACCAACAGAGTGCAGGCTAGCTCGCATCGCTGCTGCGCAAGAACTTC 543  
QY 152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg 171  
DB 544 GATCAGGAGCTCTGGGCAAGAGGTGCCCTGACCATGAGGTCTCTCGTGCATGCGCGC 603  
QY 172 ProProGluGlyIleProProAlaGluValClnuTrpLeuArgAsnGluAspLeuValAsp 191  
DB 604 CGCGCGAGGGGGTCTGTGGCGGAGGTGAATGGCTCAAGATGAGGATGTCATCGAC 663  
QY 192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211  
DB 664 CCACCCAGGACACCAACTTCTCTGCTCACCATCGACCAACCTCATCTCGCGCAGGCC 723  
QY 212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArg 231  
DB 724 CGCTGTGCGACACTGCGCAACTATACCTGCGTGCGCAAGAACATCGTGGCCAAACGCGCG 783  
QY 232 SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrp 251  
DB 784 AGCACCACTGCCACCGTCATCTGTACGTGATGGCGGTGTCTCAGTGGCGAGAGTGG 843  
QY 252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271  
DB 844 TCACCTGTCTCCAAACCGCTGTGGCGGAGGTGCGAGAGCGCACCCGACCTGCACCAAC 903  
QY 272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlnAsnValHisAspArgThrVal 291  
DB 904 CCCGCTCCACTCAACGAGGGGCTTCTGCGAGGGCCAG---GCATTCCAGAAACCGCC 960  
QY 292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys 311  
DB 961 TGCACCACTCTGCCAGTCGATGGGCGGTGGAGGAGTGAGCAAGTGTGAGCTGC 1020  
QY 312 GlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGly 331  
DB 1021 AGCACTGAGTGTGCCACTGGCGTAGCGGCGAGTGTGCGGCCCGCCACCCCAAGACGGA 1080  
QY 332 GlyGluClnCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351  
DB 1081 GGCGGTGACTGCACGGGAGCGCTGCTGACTCTAAGAACTGCACAGATGGGCTGTGCATG 1140  
QY 352 HisSer-----AlaSerGlyProGluAspValAlaLeuTyr 363  
DB 1141 CAAAGTGAGCTGTCCCGGAGTGTCTGGAGGCTCTGAGG-----GATGCGGCGCTGTAT 1194  
QY 364 ValGlyLeu-----IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382  
DB 1195 GCGGGGTCTGGTGGGCGCATCTTCGTGTGTCGAATCTCTCATGGCGGTGGGGGTGGT 1254  
QY 383 ValTyrCysArgLysLysGluGlyLeuAspSerValAlaAspSerSer---IleLeu 401  
DB 1255 GTGTACCGCCGCAACTGCGGTGACTTTCGACACAGACATCACTGACTCATCTGCTGCGCCTG 1314

QY 402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421  
DB 1315 ACTGTGTTTCCACCCGCTCACTTAAACGCGAGGCCAGTAACCCGAGCTCTTA 1374  
QY 422 -----ThrIleGlnProAspLeuSerThrThr-----ThrTyrGlnGlySerLeuCys 438  
DB 1375 CACCCCTCTGTGCTCTGACCTGACAGCCAGCGCGGCATCTACCGCGGACCGGTAT 1434  
QY 439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458  
DB 1435 GCCTGCGAGACTCC---ACCGACAAATCCCATGACCACTCTCTCTGCTGGACCC 1491  
QY 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472  
DB 1492 TTACCCAGCCCTTAAGTCAAGTCAAGTCTACAGCTCCAGCACACGCGCTCTGGGCCAGGCTG 1551  
QY 473 SerGluAlaGluPheValSerArgLeuSerThrGlnAsnTyr----- 487  
DB 1552 GCAGATGGGCTGACCTGTCTGGGGGTCTTGGCCCTGGCACATACCCTAGCGATTTCGCC 1611  
QY 488 -----PheArgSer----- 490  
DB 1612 CGGACACCCACTCTCTGCGACCTGCGCGAGCGCGAGCTCGGTCCCGAGAGCTCTTGGGC 1671  
QY 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510  
DB 1672 CTGCCCCGAGACCCAGGAGCGGTGCGCGACCTTTGGCTGCTGGGTGGGAGGCTC 1731  
QY 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530  
DB 1732 AGCATCCCGCGCACGAGTGTGCTGCTGCTGCCAATGAGGCAATTCGCCAGGCGAAG 1791  
QY 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549  
DB 1792 TTCTACGAGATGATCTACTCATCAACAGCGAGAAAGTACC---CTGCCGCTTTCAGAA 1848  
QY 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569  
DB 1849 GGGACCCAGACAGATTGAGCCCTCGTGTGACCTGTGGACCCACAGGCGCTCTGCTGTC 1908  
QY 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589  
DB 1909 CGCCCGCTCATCTCCACTGCGCCACTGTGCGAAGTCACTGCGCGCTGAGTGTGATCTTT 1968  
QY 590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609  
DB 1969 CAGCTCAAGACCCAGGCGCCACAGGGCCACTGGAGCAGAGGTGGTGGACCTGGATGAG 2028  
QY 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629  
DB 2029 GAGACCCCTGAACACACCCCTGCTACTGCGAGCTGGAGCCCGAGGCTGTCACTCTGCTG 2088  
QY 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649  
DB 2089 GACCAGCTGGCGCACTACGTGTTTCCAGCGGCGAGTCTATTCCCGCTCAGCAGTCAAGCG 2148  
QY 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669  
DB 2149 CTCAGCTGGCGGTCTTCGCGCCCGCCCTCTGCACTCCCTGGAGTACAGCTCCGCGTCC 2208  
QY 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689  
DB 2209 TACTGCTGGAGGACACGCTGTAGCACTCAAGAGGTGTGGAGCTGGAGCGGACTCTG 2268  
QY 690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709  
DB 2269 GCGGATATCTGTGGAGGAGCGAAGCCGCTAATGTTCAAGGACAGTATCACCAACCTG 2328  
QY 710 ArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyr 729  
DB 2329 CGCTCTCCCTCCATGACCTCCCGCCATGCCCATTTGGAGGAGCAAGCTGCTGGCCAAATAC 2388  
QY 730 GlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749









QY 356 GlyProGluaspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374  
 Db 1136 GGG-----GATGGGGCTGTATGCGGGGCTGTGGTGGCCATCTTCGTGGTGGCGCA 1189  
 QY 375 LeuLeuLeuValLeuLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp 394  
 Db 1190 ATCTCTATGGCGGTGGGGGTGGTGTACGCGCGCAACTGCCGTGACTTCGACACAGAC 1249  
 QY 395 ValAlaaspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413  
 Db 1250 ATCACTGACTCATCTGCTGCGCTGACTGGTGGTTCACCGCGTCAACTTAAACGCGCA 1309  
 QY 414 LysAlaaspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr 431  
 Db 1310 AGCCCGAGTAACCGCGAGCTCTACACCCCTCTGTGCTGTGACCTGACGACGAGCGCC 1369  
 QY 432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheHisLeu 450  
 Db 1370 GGCATCTACCGCGGACCGGTGTATGCCCTGCAGGACTCC---ACCGACAAATCCCATG 1426  
 QY 451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeuHisHisSerSer 470  
 Db 1427 ACCAACTCTCTCTGCTGGACCCCTTACCAGCCCTTAAGGTCAAGTCTACAGCTCCAGC 1486  
 QY 471 ProThr-----SerGluAlaGluGluPheValSerArgLeuSerThr 484  
 Db 1487 ACCACGGGCTGTGGCCAGCGCTGGCAGATGGGGCTGACTGCTGGGGTCTTCCCGCT 1546  
 QY 485 GlnAsnTyr-----LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502  
 Db 1547 GGCACATACCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGGCGAGCGCCAGC 1606  
 QY 491 -----LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502  
 Db 1607 CTCGGTTCGCCAGCAGCTCTTGGGGCTGCCCGAGACCGCAGGAGCAGCGTTCAGCGGCA 1666  
 QY 503 PheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522  
 Db 1667 TTTGGCTGCTGGTGGGAGGCTCAGCATCCCGCGCAGGAGGCTCAGCTTCTGCTGGTGGCC 1726  
 QY 523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu 542  
 Db 1727 AATGAGCCATTCGCCAGGCAAGTCTACAGATGTATCTACTCATCAACAGGCGAGAA 1786  
 QY 543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561  
 Db 1787 AGTACC---CTGCCGCTTTCAGAAAGGACCCAGACAGTATTGAGCCCTCGTGGTACCTGT 1843  
 QY 562 GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581  
 Db 1844 GGACCCACAGGCTCTGCTGTGCGCGCGCTATCTCTACCATGCCCTGCTGGCGAA 1903  
 QY 582 ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGlu 601  
 Db 1904 GTCACTGCCGTGACTGATCTTTCAGTCAAGACCCAGGCGCCACAGGGCCACTGGGAG 1963  
 QY 602 GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621  
 Db 1964 ---GAGGTGGTGGACCTGGATGAGGAGACCCCTGAACACACCCCTGCTACTGCGAGCTGGAG 2020  
 QY 622 AlaserAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641  
 Db 2021 CCAGGGCTGTACATCTCTGCTGGACGAGTGGGACCTACGTGTTCCAGGGGAGTCC 2080  
 QY 642 LeuSerValAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThr 661  
 Db 2081 TATTCCCGCTCAGCAGTCAAGCGGCTCAGCTGGCGGTTCGCGCCCGCCCTCTGCACC 2140  
 QY 662 SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuGlu 681  
 Db 2141 TCCCTGGAGTACAGCTCCCGGTCTACTGCTGGAGGACACGCTGTAGCACTGAAGGAG 2200

QY 682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuLeuGlnProArgValLeuHis 701  
 Db 2201 GTCTGAGCTGAGCGGAGCTCTGGCGGATACTTGTGGAGGAGCCGAACCCCTATG 2260  
 QY 702 PheLysaspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721  
 Db 2261 TTCAAGGACAGTTTACCACAACTTGGCGCTCTCCCTCCATGACCTCCCCCAATGG 2320  
 QY 722 LysSerLysLeuLeuValSerTyrGlnGluLeuProPheTyrHisIleTrpAsnGlyThr 741  
 Db 2321 AGAGCAAGCTGTGGCCAAATACAGGAGATCCCTTCTATCACATTTGGAGTGGCAGC 2380  
 QY 742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761  
 Db 2381 CAGAAGCCCTTCCACTGCATTTTCCCTGGAGGACAGCTTGGCTCCACAGAGCTC 2440  
 QY 762 AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781  
 Db 2441 ACCTGCAAGATCTCGCTGGCGCAAGTGAAGGGGAGGCGCAGATATTCAGCTGCATACC 2500  
 QY 782 AsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAlaGly--- 800  
 Db 2501 ACTTGGCA---GAGACACCTGTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGC 2557  
 QY 801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
 Db 2558 ACTGTACACCCAGCTGGGACCTTATGCCCTTCAAGATCCCACTGCTCCATCCGCCAGAG 2617  
 QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839  
 Db 2618 ATATGCAACAGCTTAGATGCCCACTCAACGCGGCAATGACTGGCGGATTTAGCACAG 2677  
 QY 840 LysLeuHisLeuaspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859  
 Db 2678 AAGCTCTTATGGACCGGTACTGTAATTTTCCACCAAGGAGCCCGCGGTGTG 2737  
 QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879  
 Db 2738 ATCTTGACCTCTGGGAGCTTGCACGAGACGATGGGACCTCAACAGCTTGGCGAGT 2797  
 QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
 Db 2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGTGTGGTGTGTGGCCACCGCGGAGTGC 2857  
 RESULT 13  
 ID AAS21316 standard; cDNA; 3884 BP.  
 XX AAS21316;  
 AC AAS21316;  
 XX 24-OCT-2001 (first entry)  
 DE Human cDNA sequence encoding for PRO4326 polypeptide.  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;  
 KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;  
 KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;  
 KW A-peptide; factor VIIA; gene therapy; ss.  
 XX Homo sapiens.  
 XX WO200140466-A2.  
 XX 07-JUN-2001.  
 XX 01-DEC-2000; 2000WO-US032678.  
 PF 01-DEC-1999; 99WO-US028301.  
 PR 01-DEC-1999; 99WO-US028634.  
 PR 02-DEC-1999; 99WO-US028551.  
 PR 02-DEC-1999; 99WO-US028564.  
 PR 02-DEC-1999; 99WO-US028565.  
 PR 09-DEC-1999; 99US-0170262P.

PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 20-DEC-1999; 99WO-US030919.  
 PR 30-DEC-1999; 99WO-US031243.  
 PR 30-DEC-1999; 99WO-US031274.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000277.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 18-FEB-2000; 2000WO-US004342.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 01-MAR-2000; 2000WO-US005601.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 03-MAR-2000; 2000US-0187202P.  
 PR 10-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 21-MAR-2000; 2000WO-US007532.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 05-JUN-2000; 2000US-0209832P.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 10-NOV-2000; 2000WO-US030873.

PA (GETH ) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

DR WPI; 2001-408281/43.  
 DR P-PSDB; AAU12244.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing PRO  
 PT polypeptides, and detect the presence of mammalian tumors e.g. lung,  
 PT breast, prostate, cervical.

XX Claim 3; Fig 145; 813pp; English.

XX AAS21244-AAS21518 encode for novel human secretory and transmembrane PRO  
 CC polypeptides. The PRO polypeptides are useful to detect other PRO  
 CC polypeptides, to link bioactive molecules to cells expressing PRO  
 CC polypeptides, to modulate biological activities of cells expressing PRO  
 CC polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumors by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample. Some  
 CC of the 275 sequences are also useful to stimulate the release of tumour  
 CC necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or  
 CC differentiation of chondrocytes, the proliferation or gene expression in  
 CC pericyte cells, the release of proteoglycans from cartilage, the  
 CC proliferation of inner ear utricular supporting cells or of T-  
 CC lymphocytes, the release of a cytokine from peripheral blood monocytes  
 CC (PBMCs), or the proliferation of endothelial cells. Some of the PRO  
 CC polypeptides may modulate glucose or free fatty acid uptake by skeletal  
 CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor  
 CC VIIA. The PRO polypeptides can be used in assays to identify molecules  
 CC involved in binding interactions. The polynucleotides encoding PRO  
 CC polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy  
 XX Sequence 3884 BP; 767 A; 1278 C; 1162 G; 677 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,96e-161 Length: 3884  
 Score: 2479.00 Matches: 493  
 Percent Similarity: 67.54% Conservative: 150  
 Query Local Similarity: 51.79% Mismatches: 249  
 Query Match: 51.79% Indels: 60  
 DB: 4 Gaps: 18

US-09-970-944-2 (1-899) x AAS21316 (1-3884)

Qy 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAaTrp--- 19  
 Db ATGGGGCCCGAGCGAGCTCGGGGCGCGTGTGTGGCTGTGTGTGGTGTGTGTGGAC 457  
 Qy 20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36  
 Db CCGAGGCTGAGCAAGCAGGCACTGATTCTGGCAGCGAGGTGTCTCCCTGACTCTCTCCG 517  
 Qy 37 GlyAlaAsnProAspLeuProHisPheLeuValGluProGluAspValTyrIleVal 56  
 Db TCAGCGCCAGCAGCGCGCTGCTTCTTCTGAGGAGCCACAGGACGCTACATTGTG 577  
 Qy 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76  
 Db AAGAACAAGCCTGTGGAGCTCGCTGCGCGCCTTCCCGCCACACAGATCTACTTCAAG 637  
 Qy 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96  
 Db TGAACGCGAGTGGGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 697  
 Qy 97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116  
 Db ACCGCGCTGGGTGCGGAGGTGAGTGCAGGTGTCGCGGAGCAGGTGGAGGAGCTC 757  
 Qy 117 PheGlyLeuGluGluTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136  
 Db TTTGGGTGAGGATTAATCTGTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 817  
 Qy 137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156  
 Db AGTCGCGCGAGCTACGTCCGATCGCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 877  
 Qy 157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProGluGluIle 176  
 Db GGCAGAGAGTGGCTGGAGCATGAGGTCTCTCTGAGTGGCGCGCGCGAGGGGTG 937  
 Qy 177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196  
 Db CCTGTGCGCGAGTGGATGGCTCAAGATGAGGATGTCATCGACCCACCAGGACACC 997  
 Qy 197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216  
 Db AACTTCTGTCTACCATCGACCAACACCTCATCTCGCCAGCGCGCGCTGTGGACACT 1057  
 Qy 217 AlaAsnTrpThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAla 236  
 Db GCCAACTATACCTGCTGGCCCAAGAACTCTGGCCAAAGCCCGGAGACCACTGCCACC 1117  
 Qy 237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256  
 Db GTCATCTGTACGTGAATGGCGCTGGTTCAGCTGGCAGAGTGGTTCACCTGCTCAAC 1177  
 Qy 257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276  
 Db CGTGTGGCGAGGCTGGCAGAGCGCACCGGACCTGCACCAACCCCGCTCCACTCAAC 1237  
 Qy 277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuVal 296  
 Db GGAGGGGCTTCTGCGAGGGCCAG---GCATTCCAGAAAGACCCGCTGCACCACTGTC 1294  
 Qy 297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316  
 Db CCAGTGCATGGGCGTGGAGGAGTGGCAAGTGTGTGAGCTGAGCAGCTGAGTGTGCC 1354

```
Qy 317 HistArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
Db 1355 CACTGGCGTAGCGGAGTGCATGGCCGCCACCCAGAACGAGCGCGTGAAGTGCAGC 1414
Qy 337 GlyThrAspLeuAspThrArgAsnGlySerAspLeuCysValHisSer----- 353
Db 1415 GGGACGCTGCTCAGACTCTAAGACTTGCACATGGGCTGTGTCATGCATAAATAAGAAACT 1474
Qy 354 -----AlaSerGlyProGluAspValAlaLeuTyr 363
Db 1475 CTAAGCGACCCCAACAGCACCTGCTGGAGGCTCAGGG-----GATGGCGGCTGTAT 1528
Qy 364 ValGlyLeu-----IleAlaValAlaValCysLeuValLeuLeuLeuLeu 382
Db 1529 GCGGGGCTCGTGGGCGCATCTTCGTGTCGTCGAATCTCTCATGGCGGTGGGGTGGTG 1588
Qy 383 ValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
Db 1589 GTGTACCGCGCAACTGCCGTGACTTCGACACACATCCTGACTCATCTGCTGCCCTG 1648
Qy 402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
Db 1649 ACTGGTGGTTCCACCCCGCTCAACTTTAAGACGGCAAGGCCAGCAACCCGCGAGCTCCTA 1708
Qy 422 -----ThrIleGlnProAspLeuSerThrThr-----ThrTyrGlnGlySerLeuCys 438
Db 1709 CACCCCTCTGGCTCTCCCTGACTGACACCGCGCGGCACTACCGCGGACCCGCTGTAT 1768
Qy 439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAenGlyHisLeuLeuSerPro 458
Db 1769 GCCCTGCAGGACTCC---ACGACAAAATCCCATGACCAACTCTCTCTGCTGGACCCC 1825
Qy 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472
Db 1826 TTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACACCGGCTCTGGGCCAGGCGCTG 1885
Qy 473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr----- 487
Db 1886 GCAGATGGGGCTGACCTGCTGGGGGTCTTGGCGCTGGCACATACCTTAGCGATTTCGCC 1945
Qy 488 -----PheArgSer----- 490
Db 1946 CGGGACACCCACTTCTCCTGACCTGGCAGCGCCAGCCTCGTTCCTCCAGCAGCTCTGGGC 2005
Qy 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510
Db 2006 CTGCCCGGAGACCCAGGAGCAGCGTCAAGCGGCACCTTTGGCTGCCTGGGTGGAGGCTC 2065
Qy 511 MetIleProAsnThrGlyIleSerLeuLeuLeuProAspAlaIleProArgGlyLys 530
Db 2066 AGCATCCCGGCACAGGGGTGAGTCTGTGTGCCCAATGGAGCCATTCCTCCAGGCAAG 2125
Qy 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549
Db 2126 TTCTACGAGATGTATCTACTCATCAACAAGCAGAAAGTACC---CTCCCGCTTTCAGAA 2182
Qy 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValLeuLeuThr 569
Db 2183 GGGACCCAGACAGATTATGAGCCCTCGGTGACCTGTGGACCCACAGGCGCTCTGCTGTGC 2242
Qy 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTyrSerLeu 589
Db 2243 CGCCCGCTCATCTTCACATGCCCTGCTGCCAAGTCACTGCGCGTGAAGTCACTGATCTTT 2302
Qy 590 ArgLeuLysLysGlnSerCysGluGlySerTyrGluGlnAspValLeuHisLeuGlyGlu 609
Db 2303 CAGCTCAAGACCCAGGCCACACAGGCCACTGGGAG---GAGGTGGTGACCTCGATGAG 2359
Qy 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629
Db 2360 GAGACCTTGAACACACACCTGCTACTGTCAGCTGGAGGCCCGAGGGCTGTCACATCTGCTG 2419
```

```
Qy 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649
Db 2420 GACCAGCTGGGCACCTACGTGTTCCACGGCGAGTCTTATTCCTCCCTCAGCAGTCAAGCGG 2479
Qy 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669
Db 2480 CTCACAGCTGGCGCTCTTCGCCCGCCCTCTGCACCTCCCTGGAGTAGCAGCTCCGGGTC 2539
Qy 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689
Db 2540 TACTGCCTGGAGGACCGCTGTAGCACCTGAAGAGGTGTGGAGCTGGAGCGGACTCTG 2599
Qy 690 GlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709
Db 2600 GCGGATACCTTGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCAACACTG 2659
Qy 710 ArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeuValSerTyr 729
Db 2660 CGCCTCTCCCTCCATGACCTCCCTCCATGCTGGAGGAGCAAGCTGCTGGCCAAATAC 2719
Qy 730 GlnGluIleProPheTyrHisIleTyrAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749
Db 2720 CAGGAGATCCCTTCTATCATATTGGAGTGGCAGCAAGGCCCTCCACTGACACTTC 2779
Qy 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTyrValTyrGln 769
Db 2780 ACCCTGGAGAGGCACACTTGGCTCCACAGAGCTCACCTGCAAGATCTGGTGGCGCAA 2839
Qy 770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789
Db 2840 GTGGAAGGGAGGGCCAGATATTCAGCTGCATACCACTCTGGCA---GAGACACCTGCT 2896
Qy 790 AlaGlnLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807
Db 2897 GGCTCCCTGGACACTCTCTGCTCTGCCCCCTGGCAGCACTGTCAACCCAGCTGGGACCT 2956
Qy 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827
Db 2957 TATGCTTCAAGATCCCACTGCTCCATCCGCGAGAGATATGCAACAGCTAGATGCCCCC 3016
Qy 828 CysArgArgGlyAlaAspTyrArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db 3017 AACTCACGGGCAATGACTGGCGATGTTAGCAGAGAAGCTCTCTATGGACCGGTACTG 3076
Qy 848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTyrGluAlaArg 867
Db 3077 AATTACTTTGCCAAAGCGAGCCCAACCGGTGTGATCTCGGACCTCTGGGAAGCTCTG 3136
Qy 868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887
Db 3137 CAGCAGCAGCATGGGACCTCACAGCTGGCGAGTGCCTTGGAGGAGATGGGCAAGAGT 3196
Qy 888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 3197 GAGATGCTGTGGTGTGGCCACCGAGCGGAGCTGC 3232
RESULT 14
ACD23925
ID ACD23925 standard; cDNA; 3884 BP.
XX
AC ACD23925;
XX
DT 26-AUG-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO4326 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; antiinflammatory;
KW antiarteriosclerotic; cardiant; anti-infertility; anti-HIV; cytostatic;
KW antidiabetic; gene therapy; tumour necrosis factor (TNF)-alpha release;
KW TNF-alpha release; cell proliferation; cell differentiation;
KW gene expression modulator; proteoglycan release; cytokine release;
KW tumour; inflammatory disease; organ failure; atherosclerosis;
KW cardiac injury; infertility; birth defect; premature aging; AIDS;
```

acquired immunodeficiency syndrome; cancer; diabetic complication;  
chromosome mapping; gene mapping; pharmaceutical; diagnostic; biosensor;  
bioreactor; tissue typing; gene; ss.

Homo sapiens.

US2003032156-A1.

13-FEB-2003.

06-MAY-2002; 2002US-00140474.

31-MAR-1997; 97WO-US005230.

12-JUN-1998; 98WO-US012456.

14-JUL-1998; 98WO-US014552.

28-AUG-1998; 98WO-US017888.

10-SEP-1998; 98WO-US018824.

14-SEP-1998; 98WO-US019093.

14-SEP-1998; 98WO-US019094.

16-SEP-1998; 98WO-US019177.

17-SEP-1998; 98WO-US019330.

07-OCT-1998; 98WO-US021141.

29-OCT-1998; 98WO-US022991.

20-NOV-1998; 98WO-US024855.

01-DEC-1998; 98WO-US025108.

05-JAN-1999; 99WO-US000106.

08-MAR-1999; 99WO-US005028.

10-MAR-1999; 99WO-US005190.

20-APR-1999; 99WO-US008615.

14-MAY-1999; 99WO-US010733.

02-JUN-1999; 99WO-US012252.

01-SEP-1999; 99WO-US020111.

08-SEP-1999; 99WO-US020594.

13-SEP-1999; 99WO-US020944.

15-SEP-1999; 99WO-US021090.

05-OCT-1999; 99WO-US021547.

29-NOV-1999; 99WO-US023089.

30-NOV-1999; 99WO-US028214.

30-NOV-1999; 99WO-US028313.

01-DEC-1999; 99WO-US028409.

01-DEC-1999; 99WO-US028301.

02-DEC-1999; 99WO-US028634.

02-DEC-1999; 99WO-US028551.

02-DEC-1999; 99WO-US028564.

02-DEC-1999; 99WO-US028565.

16-DEC-1999; 99WO-US030095.

20-DEC-1999; 99WO-US030911.  
20-DEC-1999; 99WO-US030999.  
22-DEC-1999; 99WO-US030720.  
30-DEC-1999; 99WO-US031243.  
30-DEC-1999; 99WO-US031274.  
05-JAN-2000; 2000WO-US000219.  
06-JAN-2000; 2000WO-US000277.  
06-JAN-2000; 2000WO-US000376.  
11-FEB-2000; 2000WO-US003565.  
18-FEB-2000; 2000WO-US004341.  
18-FEB-2000; 2000WO-US004342.  
22-FEB-2000; 2000WO-US004414.  
24-FEB-2000; 2000WO-US004914.  
24-FEB-2000; 2000WO-US005004.  
01-MAR-2000; 2000WO-US005601.  
02-MAR-2000; 2000WO-US005746.  
02-MAR-2000; 2000WO-US005841.  
10-MAR-2000; 2000WO-US006319.  
15-MAR-2000; 2000WO-US006884.  
20-MAR-2000; 2000WO-US007377.  
21-MAR-2000; 2000WO-US007532.  
30-MAR-2000; 2000WO-US008439.  
17-MAY-2000; 2000WO-US013705.  
22-MAY-2000; 2000WO-US014042.  
30-MAY-2000; 2000WO-US014941.

PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001US-00796498.  
PR 01-MAR-2001; 2001WO-US006520.  
PR 09-MAR-2001; 2001US-00802706.  
PR 14-MAR-2001; 2001US-00808689.  
PR 22-MAR-2001; 2001US-00816744.  
PR 05-APR-2001; 2001US-00828366.  
PR 10-MAY-2001; 2001US-00854208.  
PR 18-MAY-2001; 2001US-00854280.  
PR 18-MAY-2001; 2001US-00860216.  
PR 25-MAY-2001; 2001US-00866028.  
PR 25-MAY-2001; 2001US-00866034.  
PR 25-MAY-2001; 2001WO-US017092.  
PR 01-JUN-2001; 2001US-00872035.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 05-JUN-2001; 2001US-00874503.  
PR 14-JUN-2001; 2001US-00882636.  
PR 19-JUN-2001; 2001US-00886342.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 21-JUN-2001; 2001US-00887879.  
PR 22-JUN-2001; 2001WO-US020116.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 18-JUL-2001; 2001US-00908827.  
PR 06-AUG-2001; 2001US-00924419.  
PR 09-AUG-2001; 2001US-00927796.  
PR 16-AUG-2001; 2001US-00931836.  
PR 19-DEC-2001; 2001US-00028072.  
XX

(GETH ) GENENTECH INC.

Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
Gerlicsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
Smith V, Stewart TA, Tamas D, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-341980/32.  
P-ESDB; ABO17688.

New secreted and transmembrane PRO nucleic acids, for treating  
inflammation, organ failure, atherosclerosis, cardiac injury,  
infertility, birth defects, premature aging, acquired immunodeficiency  
syndrome (AIDS), or cancer.

Claim 2; Fig 145; 660pp; English.

The invention describes an isolated nucleic acid (I) comprising, or which  
has 80 % sequence identity to, or the full-length coding sequence of, one  
of 275 nucleotide sequences, and which encodes a corresponding  
polypeptide selected from 275 amino acid sequences, where all sequences  
are given in the specification. The polypeptide encoded by (I) is used to  
detect PRO polypeptides, link a bioactive molecule to a cell expressing a  
PRO polypeptide, modulate a biological activity of a cell, stimulate the  
release of tumour necrosis factor (TNF)-alpha from human blood, modulate  
the uptake of glucose or free fatty acid by cells, stimulate or inhibit  
the proliferation or differentiation of cells or gene expression,  
stimulate the release of proteoglycans, stimulate the release of cytokine  
from peripheral blood mononuclear cells, inhibit the binding of A-peptide  
to factor VIIA, or detect the presence of tumour in a mammal. The nucleic  
acid and polypeptide encoded by it, are useful for treating inflammatory  
diseases, organ failure, atherosclerosis, cardiac injury, infertility,  
birth defects, premature aging, acquired immunodeficiency syndrome  
(AIDS), cancer, or diabetic complications. The nucleic acid is useful as  
hybridisation probes, in chromosome and gene mapping, and in generating



CC antisense RNA or DNA. The polypeptides are useful as pharmaceuticals,  
CC diagnostics, biosensors or bioreactors. Both are useful in tissue typing.  
CC This sequence encodes a novel human secreted and transmembrane PRO  
CC polypeptide

XX Sequence 3884 BP; 767 A; 1278 C; 1162 G; 677 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	2,966-161	Length:	3884
Score:	2479.00	Matches:	493
Percent Similarity:	67.54%	Conservative:	150
Best Local Similarity:	51.73%	Mismatches:	249
Query Match:	51.73%	Indels:	60
DB:	7	Gaps:	18

US-09-970-944-2 (1-899) x ACD23925 (1-3884)

```
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaLaTrp--- 19
DB ATGGGGGCGCGAGCGGAGCTGGGGCGCGCTGCTGGCACTGCTGCTGGGAC 457
QY 20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
DB CGAGGCTGACCCAGCAGGCACTGATTCTGGCAGCGAGGTCTCCCTCACTCTCCG 517
QY 37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrIleVal 56
DB TCGAGCCGACAGAGCGGCTGCGCTACTTCTTGCAGGAGCCACAGAGCGCTTACATTGT 577
QY 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
DB AAGAACAAAGCCTGTGGAGTCCGCTGCGCGCGCTTCCCGCCGACACAGATCTACTTCAAG 637
QY 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
DB TGCAAGCGGAGTGGGTGACGACAGCAACACCGTACACAGGAGGCGCTGGATGAGGCC 697
QY 97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluVal 116
DB ACCGCGCTGCGGGTGGCGAGGTGAGATCGAGGTGTCGGCGAGCGGTGGAGGAGCTC 757
QY 117 PheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrLys 136
DB TTTGGCTGGAGATTACTGGTGCCAGTCGCTGGCTGGAGTCCGAGGACACCAAG 817
QY 137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
DB AGTCGCCGAGCTAGCTCCGATCGCTTACCTGCGCAAGAACTTCGATCAGGAGCCTCTG 877
QY 157 AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIle 176
DB GGCAAGGAGGTGCGCTGACCATGAGTTCCTCTGCACTGCGCGCGGAGGGGTG 937
QY 177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspSerSerLeuAspPro 196
DB CTTGTGGCCGAGTGGATGGCTCAAGATGAGGATGTATCGATCCCAACCCAGACACC 997
QY 197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
DB AACTTCTCTGCTCACCATCGACCAACCTCATCATCGCGCAGCGCGCTGTCGACACT 1057
QY 217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaLa 236
DB GCCAACTATACCTGCTGGCCAGAACATCGTGGCCAAACGCCGAGGACCACTGCCACC 1117
QY 237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
DB GTCATCGTCTACGTGAATGGCGGTGTCAGCTGGGAGAGTGGTCCACCTGCTCCAAC 1177
QY 257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
DB CGCTGTGGCGGAGTGGAGAGCGCACCCGAGCCTGCACCAACCCCGCTCCTCAAC 1237
```

```
QY 277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
DB GGAGGGGCTTCTGCGAGGGCCAG---GCATTCCAGAAGACCGCTGACACCACTATGCG 1294
QY 297 SerValAspGlySerTrpSerProTrpSerLysTyrSerAlaCysGlyLeuAspCysThr 316
DB CCAGTTCGATGGGGCTGGACGGAGTGGAGCAAGTGTTCAGCTGCAGCACTGAGTGTGCC 1354
QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
DB CACTGGCGTAGCGCGAGTGCATGGCGCCGCCACCCAGAACGGAGGCGGTGACTGCAGC 1414
QY 337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSer----- 353
DB GGGACGCTGCTCGACTCTAAGAACTGCACAGATGGGCTGTGCATCAAAATAAGAAACT 1474
QY 354 -----AlaSerGlyProGluAspValAlaLeuTyr 363
DB CTAAGGACCCACACAGCCACTGCTGGAGGCTCAGGG-----GATGCGGCGTGTAT 1528
QY 364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382
DB GCGGGCTCGTGTGGCCATCTTCGTGTCGTGCAATCCTCATGGCGTGGGGTGGTG 1588
QY 383 ValTyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerSer---IleLeu 401
DB GTGTACCCGCCCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTCCCTG 1648
QY 402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421
DB ACTGGTGTTCACCCCGTCACTTTAAGACGGCAAGGCCCAACCCGAGCTCTTA 1708
QY 422 -----ThrIleGlnProAspLeuSerThrThr---ThrTyrGlnGlySerLeuCys 438
DB CACCCTCTGTGCTCTGCTGACCTGACAGCGCGCGCATCTACCGGACCCGTGTAT 1768
QY 439 ProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458
DB GCCCTGCGAGGACTCC---ACCGACAAATCCCATGACCAACTCTCTCTGTGGACCCC 1825
QY 459 LeuGlyGlyArgHisThrLeuHisSerSerProThr----- 472
DB TTACCCAGCCTTAAGTCAAGTCTACAGTCCACACACCGGCTGCTGGCCAGGCGCTG 1885
QY 473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr----- 487
DB SCAGATGGGGTGAACCTGCTGGGGTCTTTCGCGCTGGCACATACCTAGCATTTGCGC 1945
QY 488 -----PheArgSer----- 490
DB CGGGACACCCACTTCTGCACTGCGCAGCGCGCACGCTCGGTTCAGAGAGCTTTGGGC 2005
QY 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyArgLeu 510
DB CTGCCCCGAGCACCAGGAGCAGCGCTCAGCGCACCTTTGGCTGCTGGTGGAGGCTC 2065
QY 511 MetIleProAsnThrGlyIleSerLeuLeuLeuProProAspAlaIleProArgLys 530
DB AGCATCCCGGCACAGGGGTCAAGCTTGTGTGGCCAAATGGAGCCATTTCCCGAGGCAAG 2125
QY 531 IleTyrGluIleTyrLeuThrHisLysProGluAspValArgLeuProLeuAla--- 549
DB TTCTACGAGAGTATCTACTCATCAACAGGACGAGAAAGTACC---CTCCGCTTTCAGAA 2182
QY 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569
DB GGGACCCAGACAGTATTGAGCCCTCGGTGACCTTGGACCCACAGGCGCTCTGCTGTGC 2242
QY 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeu 589
DB CGCCCGGTCTCCTCACCATGCCCGGTGTCGGAAGTCAGTGGCCGCTGACGTGATCTTT 2302
QY 590 ArgLeuLysLysGlnSerCysGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609
```

Db 2303 CAGCTCAAGACCCAGGCCACCCAGGCGCCTGGAG--GAGTGTGACCTGGATGAG 2359  
 QY 610 GUAAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629  
 Db 2360 GAGACCTCAACACACCTGCTACTGCTCCAGCTGGAGCCAGGCGCTGCATCTCTGCTG 2419  
 QY 630 GUAAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 649  
 Db 2420 GAGACCTCAACACACCTGCTACTGCTCCAGCTGGAGCCAGGCGCTGCATCTCTGCTG 2479  
 QY 650 GUAAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 669  
 Db 2480 CTCCAGCTGGCGCTCTTCGCCCGCGCTCTGACCTCCCTGGAGTACAGCCTCGGGTC 2539  
 QY 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689  
 Db 2540 TACTGCTGGAGGACACGCTGTAGCACTGAAGGAGGTCTGGAGCTGGAGCGGACTCTG 2599  
 QY 690 GlyGlyGlnLeuLeuGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709  
 Db 2600 GCGGATCTTGGTGGAGGAGCGGAACCGCTTAATGTTCAAGACAGTTACACACACCTG 2659  
 QY 710 ArgLeuSerHisAspValProSerSerLeuTyrLysSerLysLeuValSerTyr 729  
 Db 2660 CGCCTCTCCCTCATGACCTCCCGCATGCCATTTGGAGGAGCAAGCTGTGGCCAAATAC 2719  
 QY 730 GlnGluLeuProPheTyrHisLeuTyrAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749  
 Db 2720 CAGGAGATCCCTTCTATCACATTTGGAGTGGAGCGCAGAGAGCCCTCCACATGCACTTC 2779  
 QY 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTyrValTrpGln 769  
 Db 2780 ACCCTGGAGAGGACAGCTTGGCTCCAGAGCTCCTGCAAGATCTGGCTGGCGCAA 2839  
 QY 770 ValGluGlyAspGlyGlnSerPheSerLeuPheAsnIleThrLysAspThrArgPhe 789  
 Db 2840 GTGGAAGGAGGCGGCAGATATTCAGCTGCATACCACTCTGGCA--GAGACACCTGCT 2896  
 QY 790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807  
 Db 2897 GCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCACTGTCAACCCAGCTGGGACCT 2956  
 QY 808 SerAlaPheLysIleProPheLeuLeuArgGlnLysIleIleSerSerLeuAspProPro 827  
 Db 2957 TATGCTTCAAGATCCCACTGCTCCATCCCGCAGAGATATGCAACAGCTAGATGCCCCC 3016  
 QY 828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847  
 Db 3017 AACTCAGCGGGCAATGACTGGCGGATGTTAGCACAGAAGCTCTCTATGACCGGTACCTG 3076  
 QY 848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867  
 Db 3077 AATTACTTGGCCAAAGCGAGCGCCACCGGTGTGATCTCTGGAGCTCTGCGAAGCTCTG 3136  
 QY 868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887  
 Db 3137 CAGCAGGAGCATGGGACCTCAACAGCTGGCGAGTGTCTGGAGGATGGGCAAGT 3196  
 QY 888 AspaAlaGlyLeuPheThrValSerGluAlaGluCys 899  
 Db 3197 GAGATGCTGGTGGCTGTGGCCACCGGAGGAGTGTG 3232  
 RESULT 15  
 ID ACA67066 standard; cDNA; 3884 BP.  
 AC ACA67066;  
 AC ACA67066;  
 DT 23-JUN-2003 (first entry)  
 XX cDNA encoding human PRO polypeptide #73.  
 DE  
 XX

KW Human; PRO polypeptide; secreted and transmembrane protein;  
 KW anti-PRO antibody; diagnostic assay; gene expression; diabetes;  
 KW bone disorder; cartilage disorder; rheumatoid arthritis; obesity;  
 KW sports injury; osteoarthritis; hyper-insulinaemia; hypo-insulinaemia;  
 KW hearing loss; coagulation disorder; stroke; heart attack; cardiometabolic;  
 KW antidiabetic; anorectic; vulnary; antiarthritic; osteopathic;  
 KW antirheumatic; auditory; cerebroprotective; angiogenic; gene; ss.  
 XX Homo sapiens.  
 OS US2003004311-A1.  
 PN 02-JAN-2003.  
 XX 19-DEC-2001; 2001US-00028072.  
 XX 18-JUN-1997; 97US-0049911P.  
 PR 26-AUG-1997; 97US-0056974P.  
 PR 17-SEP-1997; 97US-0059113P.  
 PR 17-SEP-1997; 97US-0059115P.  
 PR 17-SEP-1997; 97US-0059117P.  
 PR 17-SEP-1997; 97US-0059122P.  
 PR 17-SEP-1997; 97US-0059184P.  
 PR 18-SEP-1997; 97US-0059263P.  
 PR 19-SEP-1997; 97US-0059352P.  
 PR 19-SEP-1997; 97US-0059588P.  
 PR 19-SEP-1997; 97US-0059836P.  
 PR 17-OCT-1997; 97US-0062250P.  
 PR 17-OCT-1997; 97US-0062285P.  
 PR 17-OCT-1997; 97US-0062287P.  
 PR 17-OCT-1997; 97US-0063755P.  
 PR 24-OCT-1997; 97US-0062814P.  
 PR 24-OCT-1997; 97US-0062816P.  
 PR 24-OCT-1997; 97US-0063045P.  
 PR 24-OCT-1997; 97US-0063082P.  
 PR 27-OCT-1997; 97US-0063127P.  
 PR 27-OCT-1997; 97US-0063327P.  
 PR 27-OCT-1997; 97US-0063329P.  
 PR 28-OCT-1997; 97US-0063550P.  
 PR 28-OCT-1997; 97US-0063561P.  
 PR 29-OCT-1997; 97US-0063704P.  
 PR 29-OCT-1997; 97US-0063733P.  
 PR 29-OCT-1997; 97US-0063735P.  
 PR 29-OCT-1997; 97US-0063738P.  
 PR 03-NOV-1997; 97US-0064248P.  
 PR 12-NOV-1997; 97US-0064809P.  
 PR 12-NOV-1997; 97US-0065186P.  
 PR 21-NOV-1997; 97US-0065846P.  
 PR 24-NOV-1997; 97US-0066364P.  
 PR 24-NOV-1997; 97US-0066453P.  
 PR 24-NOV-1997; 97US-0066511P.  
 PR 24-NOV-1997; 97US-0066770P.  
 PR 11-DEC-1997; 97US-0069212P.  
 PR 11-DEC-1997; 97US-0069278P.  
 PR 16-DEC-1997; 97US-0069334P.  
 PR 16-DEC-1997; 97US-0069694P.  
 PR 23-JAN-1998; 98US-0072320P.  
 PR 04-FEB-1998; 98US-0073612P.  
 PR 09-FEB-1998; 98US-0074086P.  
 PR 12-MAR-1998; 98US-0077791P.  
 PR 20-MAR-1998; 98US-0078910P.  
 PR 25-MAR-1998; 98US-0079294P.  
 PR 27-MAR-1998; 98US-0079663P.  
 PR 31-MAR-1998; 98US-0079728P.  
 PR 12-JUN-1998; 98US-0080165P.  
 PR 14-JUL-1998; 98US-008014552.  
 PR 28-AUG-1998; 98US-008012456.  
 PR 10-SEP-1998; 98US-008017888.  
 PR 14-SEP-1998; 98US-008018824.  
 PR 14-SEP-1998; 98US-008019093.  
 PR 14-SEP-1998; 98US-008019094.  
 PR 14-SEP-1998; 98US-008019177.



QY 297 SerValAspGlySerTyrProTyrSerLysTyrSerAlaCysGlyLeuAspCysThr 316  
 Db 1295 CCAGTCGATGGCGGTGGAGGAGTGGACGAGTGGTCCAGCTGACGACTGAGTGGCC 1354  
 QY 317 HisTyrArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336  
 Db 1355 CACTGGCGTAGCGCGAGTGCATGGCGGCCCCACCCAGAACGAGCGCGTGAAGTGCAGC 1414  
 QY 337 GlyTyrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSer 353  
 Db 1415 GGGACCGCTGCTCAGCTCTAAGAACTGCACAGATGGCGTGGCATGCAAAATAAGAAACT 1474  
 QY 354 -----AlaSerGlyProGluAspValAlaLeuTyr 363  
 Db 1475 CTAAGCGACCCCAACACCCACCTGCTGGAGGCTCAGGG-----CATGGCGGCTGTAT 1528  
 QY 364 ValGlyLeu-----IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382  
 Db 1529 GCGGGGCTCGTGTGGCCATCTTCGTGGTCTGGCAATCTCATGGCGGTGGGGTGGTG 1588  
 QY 383 ValTyrCysArgGlyLysGluGlyLeuAspSerAspValAlaAspSerSer 401  
 Db 1589 GTGTACCGCGCAACTCGCGTGAATTCGACACAGACATCACTGACTCATCTGCTGCCCTG 1648  
 QY 402 ThrSerGlyPheGlnProValSerLysProSerLysAlaAspAsnProHisLeuLeu 421  
 Db 1649 ACTGTGGTTTCCACCCGCTCACTTTAGACGGCAAGGCCCAACCCGACGCTCTTA 1708  
 QY 422 -----ThrIleGlnProAspLeuSerThrThr-----ThrTyrGlnGlySerLeuCys 438  
 Db 1709 CACCCCTCTGTGCTCTACCTGACGACGCGCGCGCATCTACCGCGGACCCGNGTAT 1768  
 QY 439 ProArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458  
 Db 1769 GGCCTGCGAGACTCC-----ACCGACAAATCCCGATACCACTCTCTGCTGGACCCC 1825  
 QY 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472  
 Db 1826 TTACCCAGCCTTAAAGTCAAGTCTACAGCTCCAGCACACCGGGCTCTGGGCCAGGCCCTG 1885  
 QY 473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr----- 487  
 Db 1886 GCAGATGGGGCTGACCTGCTGGGGGTCTTGGCGCTGGCACATACCTAGCGATTTCCGCC 1945  
 QY 488 -----PheArgSer----- 490  
 Db 1946 CGGGACACCACTTCTGCACTGCGCAGCGCGAGCCTCGGTTCACGAGCTCTTGGGC 2005  
 QY 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510  
 Db 2006 CTGCCCCGAGAGCCAGGAGAGCGTCAAGCGACCTTTGGCTGCTGGGTGGAGGCTC 2065  
 QY 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530  
 Db 2066 AGCATCCCCGGGCACAGGGGTGACGTGTGTGGTGGCCCAATGGAGCCATCCCCAGGGCAAG 2125  
 QY 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla 549  
 Db 2126 TTCTACGAGATGATCTACTCATCAACAAGCAGAAAGTACC-----CTCCCGCTTTTCAGAA 2182  
 QY 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValLeuLeuThr 569  
 Db 2183 GGGACCCAGACGATATTGAGCCCTTCGGTGACCTGTGGAGCCACAGGCCCTCTGCTGTGTC 2242  
 QY 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTyrSerLeu 589  
 Db 2243 CGCCCGCTCATCTCACCATGCCCACTGTGCCGAAGTCAAGTCCGCTGACTGGATCTTT 2302  
 QY 590 ArgLeuLeuLysGlnSerCysGluGlySerTyrGluGlnAspValLeuHisLeuGlyGlu 609  
 Db 2303 CAGCTCAAGACACCGGCCACCGAGGCCCTGGAG--GAGGTGTGACCCCTGGATGAG 2359  
 QY 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuLeuGluAlaSerAlaCysTyrValPheThr 629

Db 2360 GAGACCTCGAACACACACCCTGCTACTCCAGCTGGAGCCCGAGGCCCTGTCATCTCTGCTG 2419  
 QY 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649  
 Db 2420 GACCAGCTGGGCACCTACCTGTGTTCACGGCGAGTCTATTCCCGCTCAGCAGTCAAGCGG 2479  
 QY 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgVal 669  
 Db 2480 CTCACGCTGGCGCTCTTCCGCCCGCCCTCTGCACCTCCCTGGAGTACAGCCTCCGGTC 2539  
 QY 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeu 689  
 Db 2540 TACTGCTGGAGACACGCTGTAGCATTGAAGAGGTCTGGAGCTGGAGCGGACTCTG 2599  
 QY 690 GlyGlyGlnLeuLeuLeuGlnProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709  
 Db 2600 GCGGATACCTTGGTGGAGGAGCGAAACCGCTAACTGTTCAAGACAGTATTACCAACCTG 2659  
 QY 710 ArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeuLeuValSerTyr 729  
 Db 2660 CGCCTCTCCCTCCATGACCTCCCTCCATTCGAGGAGCAAGCTGCTGGCCAAATAC 2719  
 QY 730 GlnGluIleProPheTyrHisIleTyrAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749  
 Db 2720 CAGGAGATCCCTCTTATCAGATTTGGAGTGGCAGCCAGAGGCCCTCCACGTGACTTTC 2779  
 QY 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTyrValTyrGln 769  
 Db 2780 ACCCTGGAGAGCACAGCTTGGCTCCACAGAGCTCACCTGCAAGATCTGCTGGGCAA 2839  
 QY 770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789  
 Db 2840 GTGGAAGGAGGAGGCGCAGATATTCCAGCTGCATACCACTCTGGCA--GAGACACCTGCT 2896  
 QY 790 AlaGluLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807  
 Db 2897 GGCTCCCTGGACACTCTCTGCTCTGCTGCTGGCAGCAGCTGTCCACCCAGCTGGAGCT 2956  
 QY 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827  
 Db 2957 TATGCTTCAAGATCCCACTGCTCCATCCGCCAGAGATATGCACAGCCTAGATGCCCCC 3016  
 QY 828 CysArgArgGlyAlaAspTyrPargThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847  
 Db 3017 AACTCAGCGGGCAATGACTGGCGGATGTAGCACAGAAAGCTCTCTATGACCGTACTCTG 3076  
 QY 848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTyrGluAlaArg 867  
 Db 3077 AATTACTTTGCCACCAAGCGAGCCCGGCTGGTGTATCTCTGGACCTCTGGGAAGCTCTG 3136  
 QY 868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887  
 Db 3137 CAGCAGGACGATGGGGACTCAACAGCTGGCGAGTGTCTTGGAGGAGATGGCAAGAGT 3196  
 QY 888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
 Db 3197 GAGATGCTGGTGGTGTGGCCACCGACCGGAGCTGC 3232

Search completed: October 6, 2004, 18:38:43  
 Job time : 971 secs

***This Page Blank (uspio)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 6, 2004, 18:19:39 ; Search time 8838 Seconds  
(without alignments)  
4408.847 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGLPALLGIVLAAML.....AVAGLQPDAGLFTVSEAC 899

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-O=/cn2.1/USPTO spool/US09970944/runat\_05102004.112010.1626/app.query.fasta\_1.1095  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09970944 @CN 1.1 5932 @runat\_05102004.112010.1626 -NCPV=6 -ICPV=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4787	100.0	2881	6	AX527916 Sequence
2	4598.5	98.2	2752	6	AX449572 Sequence
3	4688.5	97.9	2697	6	AX451652 Sequence
4	4592.5	95.9	3992	10	MMU487852
5	4545.5	95.0	2697	6	AX268596 Sequence
6	4545.5	95.0	2697	10	RNU87305
7	4545.5	95.0	3014	6	BD057524 Netrin re
8	4379.5	91.5	3580	6	AX367094 Sequence
9	4278.5	89.4	3844	10	BC058084
10	3024.5	63.2	4294	10	AK122575
11	2819	58.9	2688	9	BC009333
12	2705.5	56.5	9299	10	MMU72634
13	2694	56.3	1787	6	BD057525
14	2688.5	56.2	9328	10	AB118026
15	2681	56.0	200968	2	AC123700
16	2679.5	56.0	2962	5	AY187310
17	2673.5	55.8	3646	9	AF055634
18	2570	53.7	2832	5	AY099459
19	2513.5	52.5	166534	2	AC034209
20	2505.5	52.3	3672	10	BC057560
21	2505.5	52.3	3672	10	BC057560
22	2499	52.2	3788	10	MMU487853
23	2497	52.2	2838	10	RNU87306
24	2490.5	52.0	2995	6	AX497288
25	2486.5	51.9	3770	9	AY126437
26	2484	51.9	2860	6	AX686445
27	2484	51.9	2860	6	AX686447
28	2479	51.8	3884	6	AX464012
29	2479	51.8	3884	9	AY358351
30	2479	51.8	3933	9	AB128132
31	2479	51.8	3935	9	AB096256
32	2472	51.6	2895	6	AX512381
33	2463	51.5	2831	6	BD057526
34	2444.5	51.1	168168	2	AC012283
35	2381	49.7	121282	9	AC027318
36	2276.5	47.6	9700	6	AX054976
37	2194.5	45.8	231407	2	AC139592
38	2148	44.9	2868	6	AX800719
39	2146	44.8	3252	9	AY358147
40	2144.5	44.8	3925	10	MMU487854
41	2140.5	44.7	2661	6	AX800717
42	2136	44.6	7002	6	BD162193
43	2136	44.6	7002	9	AB055056
44	1646	34.4	6495	9	AB075856
45	1646	34.4	6522	6	BD186447

# ALIGNMENTS

RESULT 1

AX527916  
LOCUS AX527916 2881 bp DNA linear PAT 21-NOV-2002  
SEQUENCE 1 from Patent WO0229038.  
ACCESSION AX527916  
VERSION AX527916.1 GI:25172359  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Herrmann, J.L., Rastelli, L. and Shinkets, R.A.  
TITLE Novel proteins and nucleic acids encoding same and antibodies  
directed against these proteins  
JOURNAL Patent: WO 0229038-A 1 11-APR-2002;  
Curagen Corporation (US)  
FEATURES  
source  
1..2881  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.13e-295 Length: 2881  
Score: 4787.00 Matches: 899  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-970-944-2 (1-899) x AX527916 (1-2881)  
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaLeuTrpLeu 20  
Db 87 ATGCGCGTCCGGCCGCGCTGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCGCTTGCTC 146  
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
Db 147 CGCGGCTCGGGTGCCCGACGAGTGCACCGTGGCCCAACAGTGGCTGTGGTCAACCCG 206  
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTrileValLysAsnLysPro 60  
Db 207 GACTGTCTTCCCACTTCTGTGGAGCCCGAGAGTGTACATCTGTCAAGAACAGCCCA 266  
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
Db 267 GTGCTGTGTGTCAAGCGCGTCCCGCCACGCGAGATCTTCTCAAGTGCACCGGGAG 325  
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
Db 327 TGGGTGGCGCCAGGTGGACCGATGATCGAGCGACGACAGCGGGAGCAGTGGTGGCCG 386  
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
Db 387 ACCATGAGGTCCCGATTATGCTTCAAGCAGCAGAGTTCGGAAGGTGTTCCGGCTGGAG 446  
QY 121 GluTrpTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140  
Db 447 GAATACTGGTCCAGTGGCTGGCATGGAGTCTCTCGGGCACCAAGAGTGCAGAGGCC 506  
QY 141 TyrtileArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
Db 507 TACATCCGATAGCCAGATTCGGAAGAACTTCAGCAGAGGAGCCGCTGGCCAGAGAGGTG 566  
QY 161 SerLeuGluGlnGlyLeValLeuProCysArgProProGluGlyLeProProAlaGlu 180  
Db 567 TCCCTGGAGCAGGGCATCGTGTGCTGCTGCTGCTCCACCGAGGGGCATCCCTCCAGCCGAG 626  
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
Db 627 GTGGAGTGGCTCCGGAACGAGGACCTGTGTGGACCGCTCCCTGGGACCCCAATGTATACATC 686

201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
Db 687 ACCGGGAGACACAGCTGTGTGTGGACAGGCCCGCTTGTCTGACAGGCCCACTACACC 746  
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
Db 747 TGGCTGGCCAAAGAACATCGTGGCAGCTCGCGCAGCGCTCCCGCTGTCTCATCGTCTAC 806  
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
Db 807 GTGAACCGGTGGGTGGTGGACCGAGTGTGTGTGGTGTGGACGCCAGCTGTGGGCGC 866  
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
Db 867 GGTGTGCGAGAACGAGGAGCGGAGCTGCACCAACCGCGGCTCTCAACGGGGCGCTTTC 926  
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
Db 927 TGTGAGGGGAGAAATGTCCATGACCGACCGTCTCTCTCTGTCTGTGTGGACGGC 986  
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
Db 987 AGCTGAGCGCGTGGAGCAAGTGTGGCGCTGTGGGCTGAGCTGCACCCACTGGCGGAGC 1046  
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
Db 1047 CGTGTGTGTCTGTACCCAGCACCCCGACGAGGGGAGGAGTGCACGGGCACTGACCTG 1106  
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
Db 1107 GACACCGCGCAACTGTATCCAGTGACCTCTGTGTACACAGTGTCTTGGCCCTGGAGGAGTG 1166  
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
Db 1167 GCCTCTATGTGGCTCTATCGCGCGTGGCGCTGTGGCTGTCTGTCTGTCTGTCTGTCTC 1226  
QY 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400  
Db 1227 ATCTCTGTTTATTTCGGGAGAGAGGAGGGGCTGAGCTCAGATGTGGCTGACTCGTCCATT 1286  
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
Db 1287 CTCACCTCAGCTTCCAGCCCGTCCAGCATCAAGCCAGCAAGCAGCAACCCCATCTG 1346  
QY 421 LeuThrIleGlnProAspLeuSerThrThrThrThrThrGlnGlySerLeuCysProArg 440  
Db 1347 CTCACCATCCAGCGGACCTCAGACCCACCCACCTACCAGGGCAGTCTCTGTGCCCGG 1406  
QY 441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460  
Db 1407 CAGGATGGGCCCGCCAGCCCAAGTTCAGCTCACCAGTGGCACCTGTCTCAGCCCTGGGT 1466  
QY 461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSer 480  
Db 1467 GGGCGCGCCACACACTGTGCCACCACTCTCCACCTCTCAGGGCCGAGGAGTTCGTCTCC 1526  
QY 481 ArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyr 500  
Db 1527 CGCTCTCCACCCAGAACTACTTCCGCTCTCCCGCGAGCCACCAACATGACCTAT 1586  
QY 501 GlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu 520  
Db 1587 GGGACCTTCAACTTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGGCTCCTC 1646  
QY 521 IleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLys 540  
Db 1647 ATCCCCCAGATGCATACCCCGGAGGAGATCTATGAGATCTACCTCAGCTGCACCAAG 1706  
QY 541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560  
Db 1707 CCGGAACAGGTGAGTTGCCCTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1766  
QY 561 CysGlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580



```
Db 1767 TGTGGACCCCTGGCGTCTGCTCACCGGCGAGTCATCCTGGCTATGACCACTGTGGG 1826
Qy 581 GluProSerProAspSerTrpSerLeuArgLeuLysGlnSerCysGluGlySerTrp 600
Db 1827 GAGCCAGCCCTGACAGCTGGAGCCTGGCGCTCAAAAAGCAGTCGTGTGGAGGCGAGCTGG 1886
Qy 601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeu 620
Db 1887 GAGCAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCGAGCTG 1946
Qy 621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
Db 1947 GAGGCCAGTGTGCTGCTAGCTTTCACCGAGCAGCTGGCGCCCTTTCCTGCTGGTGGAGAG 2006
Qy 641 AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys 660
Db 2007 GCCCTCAGCTGGCTGGCGCCAGCGCTCAAGCTGCTTCTGTGTGGCGCGGTGGCCCTGC 2066
Qy 661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
Db 2067 ACCTCCCTCGAGTCAACATCCGGTCTACTGCTGCTGATGACACCCAGGATGCACTCAAG 2126
Qy 681 GluValValGlnLeuLysGlnLeuGlyGlyGlnLeuLeuGlnGluProArgValLeu 700
Db 2127 GAGTGTGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTCTG 2186
Qy 701 HisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
Db 2187 CACTTCAGAGACAGTTACCACTCGGCTATCCATCCAGCATGTGCCAGCTCCCTG 2246
Qy 721 TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly 740
Db 2247 TCGAAGAGTAAAGCTCCTGTGACCTACCAGGAGATCCCTTTATCATCATCTGGAATGGC 2306
Qy 741 ThrClnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
Db 2307 ACGCAGCGGTACTTGCACCTTCCAGCTGGAGCGTGTCCAGCCCGAGCACTAGTGAC 2366
Qy 761 LeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIleAsn 780
Db 2367 CTGGCTGCAAGCTGTGGGTGTGGAGTGGAGGGCGAGCGGAGAGCTTTCAGCATCAAC 2426
Qy 781 PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly 800
Db 2427 TTCAACATCACCAAGGACACAAGGTTGCTGAGCTGCTGCTGTGGAGAGTGAAGCGGG 2486
Qy 801 ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle 820
Db 2487 GTCCCAAGCCCTGGTGGGCCCGAGTGCCTTCAAGATCCCTTCTCATTCGCCAAGAGATA 2546
Qy 821 IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys 840
Db 2547 ATTTCCAGCTGGACCAACCTGTAGCGGGGTGGCGACTTGGCGGAGCTTGGGCCCAAGAA 2606
Qy 841 LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIle 860
Db 2607 CTCCACCTGGACAGCCATCTCAGCTTCTTGTGCTCCAAAGCCAGCCAGCCAGCCATGATC 2666
Qy 861 LeuAsnLeuTrpClnAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 880
Db 2667 CTCACCTGTGGGAGGCGGCGACTTCCCAACGGCAACCTCAGCCAGCTGGCTGGCGACA 2726
Qy 881 ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2727 GTGGCTGGAGTGGGCGAGCCAGCAGCTGGCCCTTTCACAGTGTGGAGGCTGAGTGC 2783
```

RESULT 2  
AX449572  
LOCUS  
DEFINITION  
SEQUENCE 1 from Patent WO0210216.  
AX449572  
ACCESSION  
AX449572.1 GI:21698195

```
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Padigaru, M., Mezes, P., Mishra, V., Burgess, C., Casman, S.,  
Grosse, W.M., Alsbrook, J.P., Iepley, D.M., Gerlach, V.L.,  
MacDougall, J.R. and Smithson, G.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0210216-A 1 07-FEB-2002;  
Curagen Corporation (US)  
FEATURES  
source 1. 2752  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol_type="unassigned DNA"  
/db_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,45e-289 Length: 2752  
Score: 4698.50 Matches: 888  
Percent Similarity: 98.89% Conservative: 2  
Best Local Similarity: 98.67% Mismatches: 7  
Query Match: 98.15% Indels: 3  
DB: Gaps: 3  
US-09-970-944-2 (1-899) x AX449572 (1-2752)  
Qy 1 MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAaTrpLeu 20  
Db 46 ATGGCGCTCGCGCGCGCGCTGTGCCAGCGCTCTGGGCGATGCTCTGGCGCTTGGCTC 105  
Qy 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
Db 106 CGCGCGCTCGGCTGCCAGCAGAGTGCACCGTGGCCACCCAGCGCTGTGGTCCACCGC 165  
Qy 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60  
Db 166 GACTCTGCTCCCGCTCTCTCTGGTGGAGCGCGAGGATGTGTACATCTCTTCAAGTCAACGCGGAG 225  
Qy 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
Db 226 GTGCTGCTGTGTGCAAGCGCGCTGCGCGCGCGCGAGATCTTCTTCAAGTCAACGCGGAG 285  
Qy 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
Db 286 TGGGTGCGCCAGGTGCACCGATGATCGAGCGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 345  
Qy 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
Db 346 ACCATGAGGTCCGCAATTAATGTCTCAAGCGAGCAGGTGCGAGAGCGAGTGTGGGCTGGAG 405  
Qy 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThryLysSerGlnLysAla 140  
Db 406 GAATATGTTGCGAGTGCCTGGCATGGAGCTCTCGGGCACCGACCAAGAGGTGTGGGCTGGAG 465  
Qy 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
Db 466 TACATCCGATGAGCAGATTGCGCAAGAACTTTCGAGCAGAGCGCGCTGCGCAAGAGGTG 525  
Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProAlaGlu 180  
Db 526 TCCCTGGAGCAGGCGATCGTGTGCTGCTCCCGTCCAGCGAGGCGATCTCCAGCGAG 585  
Qy 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
Db 586 GTGGAGTGGCTCCGAACGAGGACCTGTGGAGCCCGCTCCCTGGAGCCCGCAATGTATACATC 645  
Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrTrp 220  
Db 646 ACGCGGAGCACAGCTGTGTGGTGGAGCGCGCGCTTGTGACACCGGCCCACTACACC 705
```

221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
Db TGCCTGGCCAGAACAATCGTGGCAGCTCGCGAGCGCCCTCCGCTGCTCATCGTCTAC 765  
241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
Db GTGAACCGTGGGTGGTGCAGCTGACCGAGTGGTTCCTCTGACGCGCCAGCTGTGGCGC 825  
261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
Db GGCTGGCAGAAACGAGCGAGCTGCACCAACCGCGCGCTCTCAACGGGGCGCTTTC 885  
281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
Db TGTGAGGGGAGAAATGTC---CAGAAACACCGCTGCGCCACCCCTGTCGCCAGTAGACGC 942  
301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
Db AGCTGGACCGCTGGAGCAAGTGTGCGCTGTGGCTGTGGACTGCACCCACTGCGGAGC 1002  
321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGlyCysGlnGlyThrAspLeu 340  
Db CGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGGAGAGTGCCAGGGCACTGACCTG 1062  
341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
Db GACACCGCGCAACTGATACGAGTACCTCTGTGTACACAGTGTCTTGGCCCTGGAGAGTG 1122  
361 AlaLeuTrpValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
Db GCCCTCTATGTGGCCCTATCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182  
381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400  
Db ATCTCTGTTTATTCGCGAAGAGAGGGGCTGGACTCAGATGTGGTGTGACTGCTGCTCAT 1242  
401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
Db CTCACCTCAGCGTTCACGCGCGTCAAGCATCAAGCCCGCAGCAAGCAGCAACCCCATCTG 1302  
421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThr 439  
Db CTCACCTCAGCGCGGCTCAGCACCAACCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 1362  
440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
Db CGGCAGGATGGCCCGAGCCCAAGTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 1422  
460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479  
Db GGTGGCGCGCCACACACTGACACAGCTCTCCCACTCTGAGGGCGAGAGTTCGTC 1482  
480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
Db TCCGCGCTCTCCACCCAGAACTACTTCGCTCCCTGCGCCCGGCGCCAGCAACATGACC 1542  
500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
Db TATGGACCTTCAACTTCTCGGGGGCGGTGATGATCCCTTAATACAGGTATCAGCGCTC 1602  
520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539  
Db CTCATCCCCCAGATGCCATACCCCGAGGAGAGATCTATGAGATCTACTCTACGCTGCAC 1662  
540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
Db AAGCCGAGAGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1722  
560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
Db AGCTGTGGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1782  
580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599

1783 GGGAGCCCAAGCCCTGACAGCTGGAGCTCGCCTCAAAAGCAGCTGTCGGAGGGCAGC 1842  
QY TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTrpTyrCysGln 619  
Db TGGGAG---CATGTGTGACCTGGGCGAGAGCGCCCTCCACCTTACTACTGCTCAG 1899  
620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
Db CTGGAGCGCAGTGCCTGCTACGTCTTCCAGCAGCAGCTGGGCGCTTGGCTTGGTGGGA 1959  
640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659  
Db GAGCCCTCAGCGTGGTGGCGCAAGCGCTCAAGCTGCTTCTGTTGGCGCGTGGCC 2019  
660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
Db TGCACCTCCCTCGAGTACAACTCCCGGTCTACTGCTGCTGATGACACCCAGATGCACTC 2079  
680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699  
Db AAGAGTGTGTGAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCAGGAGCCACGGGTC 2139  
700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
Db CTGCACCTCAAGGACAGTACCACAACCTGCGCTATCCATCCACGATGTGCCAGCTCC 2199  
720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
Db CTGTGGAAGAGTAACTCTTGTGAGTACAGAGATCCCTTTTATCATCATCTGGAAT 2259  
740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
Db GGCACGAGCGGTACTTGCACCTGACCTTACCTCTGAGAGTGTGACCCCGCAGCAGT 2319  
760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779  
Db GACCTGGCTGCAAGCTGTGGTGTGGCAGGTGGAGGGCGAGCGGAGCTTCAGCATC 2379  
780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
Db AACTTCAACATCAACAAGGACACAAGTGTGCTGAGCTGTGGTCTGGAGAGTGAAGCG 2439  
800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
Db GGGGTCCAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCCCTCATTCGGCAGAG 2499  
820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839  
Db ATATTTCCAGCCTGGACCCACCTGTAGGGGGGTGCGGACTGCGGACTCTGCGCCAG 2559  
840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859  
Db AAATCCACCTGGACCGCATCTCAGCTTCTTGGCTTCCAGCCAGCCCGCCAGCCAGCATG 2619  
860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879  
Db ATCTTCAACCTGTGGAGCGCGGACCTTCCCGACCGCAACCTCAGCCAGCTGGCTGCA 2679  
880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlnCys 899  
Db CGATGGTGTGAGTGGCGCAGCCAGCAGCTGGCTTCTTCAAGTGTGCGAGGCTGAGTGC 2739

## RESULT 3

AX451652

LOCUS

DEFINITION Sequence 1 from Patent WO0233080.

ACCESSION AX451652

VERSION AX451652.1 GI:21698587

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

linear

DNA

2697 bp

PAT 03-JUL-2002

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 Koehler, R.H.  
 Regulation of human netrin binding membrane receptor unc5h-1  
 Patent: WO 0233080-A 1 25-APR-2002;  
 Bayer Aktiengesellschaft (DE)  
 Location/Qualifiers  
 1..2697  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

**REFERENCE**  
**AUTHORS**  
**TITLE**  
**JOURNAL**  
**FEATURES**  
 source

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 1.03e-288 Length: 2697  
 Score: 4688.50 Matches: 986  
 Percent Similarity: 98.78% Conservative: 3  
 Best Local Similarity: 98.44% Mismatches: 8  
 Query Match: 97.94% Indels: 3  
 DB: 6 Gaps: 3

US-09-970-944-2 (1-899) x AX451652 (1-2697)

QY	1	MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAlaTyrLeu	20
Db	1	ATGGCCCTCGGGCCGGCCCTGTGGCCAGCGCTCCTGGGCATAGTCTCGCGCTTGGCTC	60
QY	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
Db	61	CGCGGCTCGGGTGCCAGCAGAGTGCCACCGGTGGCCACCCAGTGCCTGGTGGCAACCCG	120
QY	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro	60
Db	121	GACCTGCTTCCCACTTCTGGTGAGCCCGAGATGTGTACATCGTCAAGAACCAAGCCA	180
QY	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	181	GTGCTGCTGTGTGCAAGGCGTGCCTGGCCAGCAGTCTCTTCAAGTGCACGGGGAG	240
QY	81	TyrValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	241	TGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACGCGGACGAGTGGGCTGCC	300
QY	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	301	ACCATGAGTCCGCAATTAATGCTCAAGGACGAGTCCGAGAGGTGTTCGGGCTGGAG	360
QY	121	GluTyrTrpCysGlnCysValAlaTyrSerSerSerGlyThrThrLysSerGlnLysAla	140
Db	361	GAATACTGGTGCAGTGCCTGGATGGAGTCTCTCGGGCCACCACCAAGATCAGAGGCC	420
QY	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
Db	421	TACATCGCATACCTATTTCGCAAGAACTTCGAGCAGAGCGCTGGCCAAAGAGGTG	480
QY	161	SerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIleProAlaGlu	180
Db	481	TCCTTGAGCAGGGCATCGTGTGCTGCTCCAGCGAGGGCATCCCTCCAGCGGAG	540
QY	181	ValGluTyrLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
Db	541	GTGGAGTGGCTCGGAACAGAGACCTTGGTGGACCGCTCCGACCCCAATGTATCATC	600
QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
Db	601	ACCGGGAGCAGACCTGTGGTGGCAGACGCGCGCTTGTGACACGGCCAACTACAC	660
QY	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr	240
Db	661	TGGGTGGCCAAAGAACATCGTGGACGCTCGCCGAGCGGCTTCCGCTGTATCTGCTAC	720
QY	241	ValAsnGlyClyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg	260
Db			

Db	721	GTGAACGCTGGTGTGCTGACCGTGAGCCAGTGCTCCGTCTGAGCGCCAGCTGTGGGCGC	780
QY	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyClyAlaPhe	280
Db	781	GGCTGGCAGAAACGAGCGGAGCTGCACCAACCGCGGCTCTCAACGGGGCGCTTTC	840
QY	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
Db	841	TGTAGGGGGCAGAATGTC--CAGAAAAACAGCCTGGCCGCCCTGTGCCAGTGGACGCG	897
QY	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
Db	898	AGCTGGAGCCCGTGGAGCAAGTGTGGCCCTGTGGCTGGACTGCACCATCTGGCGGAGC	957
QY	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGlnCysGlnGlyThrAspLeu	340
Db	958	CGTGAGTGTCTGACCCAGCACCAGCGGAGGAGGAGTGCAGGCGACTGACCTG	1017
QY	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360
Db	1018	GACACCCGCAACTGTACAGTGACCTCTGTGTACACACTGTCTTGGCCCTTGAGGACGTG	1077
QY	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuValLeu	380
Db	1078	GCCCTCTATGTGGCCCTCATCGCGCTGGCCGTCTGCCCTGGTCTCTGCTGCTGTGCTC	1137
QY	381	IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle	400
Db	1138	ATCTCTCTTATTGCGGAAGAGGAGGCTGACCTCAGATGTGGCTGACTCTGCTCCATT	1197
QY	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420
Db	1198	CTACCTTCAGCTTCCAGCCCGTTCAGCATCAAGCCAGCAAGCAGACAAACCCCATCTG	1257
QY	421	LeuThrIleGlnProAspLeuSer--ThrThrThrThrThrGlnGlnGlySerLeuCysPro	439
Db	1258	CTCACCATCCAGCGGACCTCAGCACCCACCCACCACTTCCACCTCTGAGGCGGAGTTCGT	1317
QY	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu	459
Db	1318	CGGAGGATGGGCGCCAGCCCAAGTTCAGCTCACCATAAGGCGCCCTGCTCAGCCCCCTG	1377
QY	460	GlyGlyClyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal	479
Db	1378	GGTGGGCGCGCCACACACTGCACACCTCCACCTCTCCACCTCTGAGGCGGAGTTCGT	1437
QY	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499
Db	1438	TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGACCAACATGACC	1497
QY	500	TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519
Db	1498	TATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGAATACGCTC	1557
QY	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis	539
Db	1558	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTACGCTGCAC	1617
QY	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559
Db	1618	AAGCGGAGAGAGCTGAGGTGTGCCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1677
QY	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
Db	1678	AGCTGTGGACCCCTGGCGTCTGCTCACCAGGCGAGTCTCTGCTGCTGCTGCTGCTGCTG	1737
QY	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysGlnSerCysGluGlySer	599
Db	1738	GGGAGAGCCAGCCCTGACAGCTGGAGCCTCGCCCTCAAAAAGCAGCTGTCGAGGCGAG	1797
QY	600	TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln	619
Db	1798	TGGAG--GATGTGCTGCACCTGGCGGAGGAGGCGCCCTCCACCTCTACTACTGCCAG	1854

620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
1855 CTGGAGCCAGTGGCTCTAGCTTCCACGAGCAGCTGGCGCTTTGGCCCTGGTGGGA 1914  
640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659  
1915 GAGGCCCTCAGCGTGGCTGGCGAGCAGCTCAAGCTGCTTCTTTGGCGCGTGGCC 1974  
660 CysThrSerLeuGluTyrAsnLysArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
1975 TGACCTCCCTCGATACACATCCGGTCTACTGCTGCTGATGACACCCAGATGCACTC 2034  
680 LysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuGlnPuArgVal 699  
2035 AAGGAGGTGGTGGAGTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCACCGGTC 2094  
700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
2095 CTGCACCTTCAAGGACAGTTACCAACCTGGCGCTATCCATCCAGATGGCCAGCTCC 2154  
720 LeuTyrLysSerLysLeuLeuValSerTyrGlnGlnLeuProPheTyrHisIleTyrAsn 739  
2155 CTGTGGAGAGTAACTCTTGTGAGCTACAGGAGATCCCTTTATCATCATCTGGAAT 2214  
740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
2215 GGCACGAGCGGTACTTGCATGCACTTCACTCCCTGGAGCGTGTGAGCCCGAGCACTAGT 2274  
760 AspLeuAlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIle 779  
2275 GACCTGGCGCTGCAAGCTGTGGGTGTGGAGTGGAGGCGGAGGAGCTTTCAGCATC 2334  
780 AsnPheAsnLysThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
2335 AACTTCAACATCACCAAGGACACAAAGTTTCTGAGCTGCTGGCTCTGGAGAGTGAAGCG 2394  
800 GlyValProAlaLeuValGlyProSerAlaPheLysLysLeuPheLeuLysArgGlnLys 819  
2395 GGGGTCCAGCCCTGGTGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAG 2454  
820 IleIleSerSerLeuAspProCysArgArgGlyAlaAspTyrArgThrLeuAlaGln 839  
2455 ATAATTTCCAGCTTGACCCACCTGTAGGCGGGTGGCGACTGCTGGCCAG 2514  
840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859  
2515 AAATCCACCTGGAGCAGCATCTCAGCTTCTTGGCTCCAGGCCAGCCACAGCCATG 2574  
860 IleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSerGluLeuAla 879  
2575 ATCCTCAACCTGTGGAGCGCGCAGCTTCCCAACGGCAACCTCAGCAGCTGGCTGCA 2634  
880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
2635 GCAGTGGCTGAGCTGGGCGCAGCAGCAGCTGGCTTCTTCACTGCTCGAGGCTGAGTGC 2694

RESULT 4  
MMU487852 3992 bp mRNA linear ROD 24-SEP-2002  
LOCUS Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).  
DEFINITION  
ACCESSION AJ487852  
VERSION AJ487852.1 GI:22035783  
KEYWORDS netrin receptor Unc5h1; Unc5h1 gene.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1  
REFERENCE  
AUTHORS Engelkamp, D.  
TITLE Cloning of three mouse Unc5 genes and their expression patterns at mid-gestation  
JOURNAL Mech. Dev. 118 (1-2), 191-197 (2002)

MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

22239710  
12351186  
2 (bases 1 to 3992)  
Engelkamp, D.  
Direct Submission  
Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for  
Brain Research, Deutschoordenstrasse 46, Frankfurt 60528, GERMANY  
Location/Qualifiers  
1. 3992  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1. 3992  
/gene="Unc5h1"  
232. 2928  
/gene="Unc5h1"  
/codon\_start=1  
/product="netrin receptor Unc5h1"  
/protein\_id="CAD32250.1"  
/db\_xref="GI:22035784"  
/db\_xref="GOA:Q8K1S4"  
/translation="MAVRPGLWPALLGIVLTAMLRGSGAQQSATVANPVGANPDLLP  
HFIVEPEVDYIVKPKVLLYKCAVPATQIEFKNGEWMVRQVHVIERSTDGSSGLPTM  
EVRINVSROQVEKVFGLLEYVQCQVAMSSSGTTKSKAYIRIAYLRNFPQELAKEV  
SLEQGLVLCRPEGIPPAEVEMLRNEDLVDPSLDPNVITREHSLVVRQARLADTAN  
YTCVAKNIVARRBSASAALVYVNGGWSWTWESVCSACSGRWKSRSDPAPRNGERC  
GGAFCEGQVQKTACATLCPVDGSWSPWKSACGLDCTHWRSECDPAPRNGERC  
RGADLPTNCTSDLCHTSSGDEVDALYIGLVAVAVCLILILLVLLVLLYCRKKGELS  
DVADSSILTSFGFVSIKPSKADNPALLTIQDLSITTTTQSSLCPRQDGPSPFQL  
SNHLLSPGLSGRHTLHSSPTSEAEFVSRSLSTQNYFRSLPRGTSNMYGTNFPFGV  
RLMPTIGSLLIPDDAI PRGKIVEIYTLHKPEDVRLPLAGCOTLLSPVSCGPPGV  
LLRFPVILAMDHGCEPSDSRLKQSCGSEGWEDVHLGSESPSHLYCOLBAGAC  
YVTEQGRFALVGEALSVATKRLRLLLPAPVACTSLEYNIRVYCLDTHDALKVY  
QLKQGGQILIEPRVLHFQKSDHNRSLSHDPSLWKSLLVSKLVSQEIPIYHWNGT  
QVVLGCTFTLERNASTDLACKWQWQVGDGOSFNINFTKDTFPAEMLALESEG  
GGPALVGFSAKIPFLIRQKITSLDPPCSRGADWRTLAQKLHLSLFFASKESPT  
AMTLINLEWARHFFPNGLQLAAAVAGLQPDAGLFTVSEAC"

## ORIGIN

## Alignment Scores:

Pred. No.: 2.19e-282 Length: 3992  
Score: 4592.50 Matches: 858  
Percent Similarity: 97.67% Conservative: 21  
Best Local Similarity: 95.33% Mismatches: 18  
Query Match: 95.94% Indels: 3  
DB: 10 Gaps: 3

US-09-970-944-2 (1-899) x MMU487852 (1-3992)

QY 1 MetAlaValArgProGlyLeuTyrProAlaLeuLeuGlyIleValLeuAlaAlaTyrLeu 20  
Db 232 ATGGCCGTCCTGGCCCGCGCTGTGGCCAGCGCTCTGGGCATAGTCTCTACCTGCTGGCTT 291  
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
Db 292 CGTGGTTCGGTCCGCCAGCAGAGTGCCACAGTGCCCAACCCAGTGCCTGTGTCACACCA 351  
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60  
Db 352 GACCTGCTGCCCACTTCTCTGTAGAGCGGAGGACGCTGTATCATTTGTCAAGAACAAAGCCC 411  
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
Db 412 GTGCTGCTGGTGTGCAAGGCTGTGCCGCCACCCAGATCTTCTTCAAGTGCAACGGGAA 471  
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
Db 472 TGGGTTCGCCAGTTCGATCAGTCATCGAACCGCAGCACTGACGGCAGCAGCGGATTCGCA 531  
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
Db 532 ACCATGGAGTCCCGGATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAG 591

QY 121 GluTyrTrpCysGlnCysValAlaValTrpSerSerGlyThrThrLysSerGlnLysAla 140  
DB 592 GAGTACTGGTCCAGTGTGGCATGGAGCTCTCAGGACACCAAGCAAGGAGGCC 651  
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
DB 652 TACATCCGGATTGCTATTTCGCAAGAACTTTGAGCAGGAGCGCTGGCCAAAGGAGTG 711  
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180  
DB 712 TCATGGAGCAAGGCAATTGTGCTACCTTGTGCGCCCGGAGGAATCCCGCCAGCTGAG 771  
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
DB 772 GTGAGTGGCTCCGAATGAGGACCTCGTGACCCCTCCCTCGACCCCAATGTGTACATC 831  
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
DB 832 ACACGGGAGCACAGCTAGTGTGGCGGAGGCGCGCTGGCGGACACTGCCAACTACACC 891  
QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
DB 892 TGCCTGGCCAGAACATCGTGGCCCGTGGCGAAGCGCTCTGGCGCGCTCATTTGTTTAT 951  
QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
DB 952 GTGAACGGTGGGTGGTGGACCGAGTGGTCCGTCTGCAGTGGCAGCTGTGGCGGT 1011  
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
DB 1012 GGTGGGAGAAACGAGCGGAGCTGCACCAACCGCGCACCTCTCAACGGGGCGCTTC 1071  
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
DB 1072 TGTGAGGGGAGAAATGTC---CAGAAACAGCGCTGGCGCACTCTGTGCCAGTGGATGGG 1128  
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
DB 1129 AGCTGGAGGCCATGGAGTAAGTGGTCAGCTGGGGCTTGACTGCACCCACTGGCGGAGC 1188  
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
DB 1189 CGGGAGTGTCTCCAGCCAGCGCCCGCAACGGAGGTGAGGAGTGGCGGGTGTGACCTG 1248  
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
DB 1249 GACACCGCAACTGTACAGTGACCTCTGCTCGACACACCTCTTCCGCGCCCGGAGGACGTG 1308  
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
DB 1309 GCTCTCTACATCGCCCTCGTCGCGGTGGCGGTGTGCTCATCTGTGCTGTGGTCCCTC 1368  
QY 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerIle 400  
DB 1369 GTTCCTCATCTACTCCGCGCAAGAGGAGGACTGGACTCAGACCGTGGCTGACTCATCATC 1428  
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
DB 1429 CTTACTCTAGGCTTCAGGCTGTACGATCAAGCCAGCAAGCAAGCAATCCCCATCTG 1488  
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439  
DB 1489 CTCACCATCAACCGGACCTCAGCACCAACACCACTTACACGAGGAGCCCTGTGTCCC 1548  
QY 440 ArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
DB 1549 CGGACGATGGACCCAGCCCAAGTTCAGCTCTCTAATGGTCACTGTGTGAGGAGCCCTGTGTCCC 1608  
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479  
DB 1609 GGCAGTGGCGGCATACGCTGCACCAACAGCTCCCCACCTCTGAGGCTGAGGACTTCGTGTC 1668

QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
DB 1669 TCCGCGCTCTCCACCCAAACTACTTTCTGCCCCGCGGTACCAAGCAACATGGCC 1728  
QY 500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
DB 1729 TATGGGACCTTCACTTCTCGGGGGCGGCTGATGATCCCTTAACACAGGAATACGCTC 1788  
QY 520 LeuIleProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539  
DB 1789 CTCATACCCCGGAGCGCATCCCGGAGAAAGATCTACGAGATCTACTCTACTCTGAC 1848  
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
DB 1849 AAGCCAGAAAGACGTGAGGTTGCCCTAGCTGGCTGTGAGACCTCTGAGTCTATCGTT 1908  
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
DB 1909 AGCTGTGGGCCCCCAGGAGTCTCTGCTCACCGCGCCAGTCACTCTTGCATCGACCTGC 1968  
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGlyGlySer 599  
DB 1969 GGGGAGGCCAGTCCCGACAGCTGGAGCTCGCCCTCAAAAAGCAGTCTCTGTGAGGGCAGC 2028  
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619  
DB 2029 TGGAG---GACGTGTGTCACCTTGGTGGAGGAGTGGCCCTCTCATCTCTACTTGCAC 2085  
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
DB 2086 CTGGAGCCCGGGCGCTGTATGCTTCCAGGAGCAGCTAGCGCCCTTGGCCCTGGTGGGA 2145  
QY 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659  
DB 2146 GAGGCCCTCAGCTGGCTGCCCAAGCGCTCAGGCTCTCTGTTGTTGCCCTGTGGCC 2205  
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
DB 2206 TGTACGTCTCCAGTACAAACATCCGAGTGTACTGCTGCACGACACCCACGATGCTCTC 2265  
QY 680 LysGluValValGlnLeuLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699  
DB 2266 AAGGAGTGTGTGAGTGTGGAGAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTC 2325  
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
DB 2326 CTGCACTTCAAGACAGTTACCAACACTTACCTGTCTGTCCATCCACGAGCTGCCAGCTCC 2385  
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
DB 2386 CTGTGAAAGAGCAAGCTCTGTGCTACCTACGAGAGATCCCTTTTACCACATCTGGAAT 2445  
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
DB 2445 GGCACCTCAGAGTATCTGCAGTGTGCACCTTACCTCGAGCGCGTCAATCCAGCAGCAGC 2505  
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779  
DB 2506 GACCTGGCTGCAAGGTGTGGGTGTGGAGGTGGAGGAGATGGACAGAGCTTCAACATC 2565  
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
DB 2566 AACTTTAAATCACTAAGGACACGAGGTTGTGAAATGCTGGCTCTGGAGAGTGAAGGG 2625  
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLeuIleProPheLeuIleArgGlnLys 819  
DB 2626 GGGGTCCCGACCCCTGGTGGGCCCCAGTGTCTTCAAGATCCCCCTTCTCTCAATCGGCAAAAG 2685  
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839  
DB 2686 ATCATTACCAGCTGGACCCACCTTGCAGCCGGGGCGGAGCTGGCGAACTCTTAGCCGAG 2745  
QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859

```

Db      2746  AAACCTTCACTGACAGCACTTAGCTTTTGGCTTCCAGCCAGCCCTACAGCATG 2805
QY      860  ILeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db      2806  ATCTCACTATGGAGGGCGGCACTTCCCAACGGCAACCTCGCCAGCTGGCGCA 2865
QY      880  AlaValalaGlyLeuGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db      2866  GCTGTGGCGGAGTGGGCCAGCCAGATGCTGGCTCTTCCCGTGTCCAGAGCGCGATGC 2925

RESULT 5
AX268596 LOCUS 2697 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 15 from Patent WO0175440.
ACCESSION AX268596
VERSION AX268596.1 GI:16541710
KEYWORDS Rattus sp.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Cochran,S.W., Paterson,G.Y., Ohashi,Y.W., Morris,B.Y. and
Pratt,J.X.
TITLE Schizophrenia related genes
JOURNAL Patent: WO 0175440-A 15 11-OCT-2001;
WELFIDE CORPORATION (JP)
FEATURES
Source 1..2697
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN
Alignment Scores:
Pred. NO.: 1.28e-279 Length: 2697
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: 6 Gaps: 3

US-09-970-944-2 (1-899) x AX268596 (1-2697)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db 1 ATGCCCGTCCGGCCCGCGCTGTGGCAGTGTCTCTGGGCATAGTCTCTCGCGCGCTGGCTT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAlaAsnProValProGlyAlaAsnPro 40
Db 61 CGTGGTTCCGGTGCCAGCAGAGTGCACGGTGGCCAAATCCAGTGCCCGGTGCCAACCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrlleValLysAsnLysPro 60
Db 121 GACTGTGCCCCACTTCTCTGGTAGCCTGAGACGTGTACATTTGTCAAGNACAAGCCG 180
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 181 GTGTGTGTGGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTTCAAGTGTCAATGGGAA 240
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db 241 TGGGTCCCGCAGGTGCATGCAATTAAGACGACGACCGACGAGCAGCGAGTGGCA 300
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 301 ACCATGGAGGTCCTCGATCAAGTATCGGACGACAGTGTAGAGAAAGTGTTCGGCTGGAG 360
QY 121 GluTyTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db 361 GAAATACGTGTCAGTGTGTGGCATGGAGCTCTCTCGGGTACCACCAAAAGTCAGAAGGCC 420

```

```

QY 141 TyrlleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 421 TACATCCGGATTCCCTATTTCGCAAGAACTTTGACGAGGACCACTGGCCCAAGGAAGTG 480
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db 481 TCACTGGAGCAAGGCATTTGACTACCTTTGTCGCCCCCCAGAGAAATCCCCCAGCTGAG 540
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrlle 200
Db 541 GTGGAGTGGCTTCCAAATAGGACCTCGTGGAGCCCTTCCCTCCATCCAAATGTGTATC 600
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrlle 220
Db 601 ACGGGGAGCAGACCTAGTCGTGCTGAGCCCGCTGGCCGACACGCGCACTACAC 660
QY 221 CysValalaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrlle 240
Db 661 TGTGTGGCCAAAGAACATCGTAGCCCGTCGCGAAGCACCTCTGCGAGCGGTCAATTGTTAT 720
QY 241 ValAsnGlyGlyTrpSerThrTrpGluTrpSerValCysSerAlaSerCysGlyArg 260
Db 721 GTGAACCGTGGTGGTGCAGCTGAGTGGTCCGTCTGCGCGCAGCTGTGGCGGT 780
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
Db 781 GGCTGGCAGAAACGGAGCGGAGCTGCACCAACCGGCACCTCTCAACGGGGCGCTTC 840
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
Db 841 TGTGAGGGGCGAAGATGTC---CAGAAAACACGCTGCGCCACTCTGTGCCAGTGGATGGG 897
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
Db 898 AGCTGGAGTTCTGGAGTAAGTGTGAGCTGTGGCTTGTACTGCACCCACTGCGCGAGC 957
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
Db 958 CGCGAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGCGGGTGTGACCTG 1017
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db 1018 GACACCCGCAACTGTACAGTGACCTCTGCTGCACACCGCTTCTTGTGCCCGAGACGTG 1077
QY 361 AlaLeuTyrlleValGlyLeuIleAlaValAlaValCysLeuValLeuLeuValLeu 380
Db 1078 GCTCTCTACATCGGCGCTGTGCTGTGCTGTGCTCTTCTTCTGTGTGCTGGCCCTT 1137
QY 381 IleLeuValTyrlleCysArgLysGlyLeuAspSerAspValAlaAspSerSerIle 400
Db 1138 GGACTCATTTACTCTCGCAAGAGGAGGGCTGGACTCCGATGTGGCGACTGTGCTCATC 1197
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1198 CTCACCTCGGCTTCCAGCCTGTGAGTATCAAGCCAGCAAGCAGACCAACCCACCTG 1257
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrlleGlnGlySerLeuCysPro 439
Db 1258 CTCACCTCGGCGAGCAGCTCAGCACCCACCTACCTACCTACCGGGCAGTCTATGTCG 1317
QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1318 AGGAGGATGACCCAGCCCAAGTTCCAGCTCTTAATGGTCACTGTCTGAGCCCACTG 1377
QY 460 GlyGlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGluLupheVal 479
Db 1378 GGGAGTGGCCCGCATACGTTGCACACAGCTACCCACTCTGAGGCTGAGGACTGTGTC 1437
QY 480 SerArgLeuSerThrGlnAsnTyrllePheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1438 TCCCGCTCTCCACCCCAAACTACTTTGTTCCCTGCCCCCGCGCACCAACATGGCC 1497

```

QY 500 TyrGlyThrPheAsnSerLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
 Db 1498 TACGGGACCTTCAACTTCTCGGGGCGGTGATGATCCCTAATACGGGATCAGCCTC 1557  
 QY 520 LeuIleProProAspAlaIleProArgGlyIleTyrGluIleTyrLeuThrIleHis 539  
 Db 1558 CTCATACCCCGGATGCGCATCCCGGAGAAAGATCTACGAGATCTACCTCACATCGAC 1617  
 QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
 Db 1618 AAGCCAGAGAGCTGAGCTTGCCTAGCTGCTGCTAGACCCCTGCTGAGTCCAGTCGTT 1677  
 QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
 Db 1678 AGCTGTGGGCCCCCAGGAGTCTGCTACCCGGCCAGTCATCTTGCATGGACCATGT 1737  
 QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599  
 Db 1738 GGAGAGCCCAAGCTGACAGCTGAGTCTGCGCCTCAAAAGACAGTCTCGGAGGCGAGT 1797  
 QY 600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619  
 Db 1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTACCTTCCCACTCTACTACTGCGAG 1854  
 QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
 Db 1855 CTGAGGCGCGGCGCTGCTATGCTTTCAGGAGCAGCTGGGCGCTTGCCTGTAGGA 1914  
 QY 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659  
 Db 1915 GAGGCCCTCAGCGGTGCGCACCAAGCGCTCAGGCTCTCTTGTGTCCGTCGCGC 1974  
 QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
 Db 1975 TGTACGTCCCTTGATACACATCCGAGTGTACTGCTTACAGCACCCAGCGCTCTC 2034  
 QY 680 LysGluValValGlnLeuGlyLysGlnLeuGlyGlnLeuLeuGlnGluProArgVal 699  
 Db 2035 AAGGAGTGTGTCAGCTGGAGAGCAGCTAGGTGGACAGCTGATCCAGGAGCTCGGCTC 2094  
 QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
 Db 2095 CTCGACTTCAAGACAGTATTACCAACCTACGCTCTCCATCCACGACGTCGCCAGCTCC 2154  
 QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
 Db 2155 CTGTGGAAGACAGCTACTTGTGCTACGACTACGAGATCCCTTTTACCACTCTGGAAC 2214  
 QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
 Db 2215 GGCACCCAGCAGTATCTGCATCTGCACCTCACCTTGGAGCGCATCAACGCCAGCACCG 2274  
 QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779  
 Db 2275 GACCTGGCTGCAAGGTGTGGGTGGAGTGGAGCGAGATGGGAGAGCTTCAACATC 2334  
 QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
 Db 2335 AACTTCAACATCACTAAGACACAAAGTTTGTGAATTTGTGCTCTGGAGATGAAGGG 2394  
 QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
 Db 2395 GGGGTCCCAAGCTGTGGGCCCCAGTGTCTTCAAGATCCCTTCTCATTCGGCAAG 2454  
 QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839  
 Db 2455 ATCATCGCAGCTGTGACCCACCTCAGCGGGGCGGCGGAGTCTAGGCCAG 2514  
 QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859  
 Db 2515 AACTTCACTGACAGCATCTTACTCTTCTTGTGCTTCAAGCCAGCCGCTCAGCCATG 2574  
 QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879

Db 2575 ATCTCAACCTATGGAGCGCGGCACTTCCCAACGCAACCTCGGCCAGCTGCACGA 2634  
 QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
 Db 2635 GCTGTGGCGGACTGGGCAACACAGATGCTGCCCTCTTACCGTGTGCGAGCCGAGGT 2694  
 RESULT 6  
 RN087305 2697 bp mRNA linear ROD 15-MAY-1997  
 LOCUS Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds.  
 DEFINITION U87305  
 ACCESSION U87305.1 GI:2055391  
 VERSION U87305.1  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE Rattus norvegicus  
 ORGANISM Rattus norvegicus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 2697)  
 AUTHORS Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S.L. and  
 Tessier-Lavigne, M.  
 TITLE Vertebrate homologues of C. elegans UNC-5 are candidate netrin  
 receptors  
 JOURNAL Nature 386 (6627), 833-838 (1997)  
 MEDLINE 97271897  
 PUBMED 9126742  
 REFERENCE 2 (bases 1 to 2697)  
 AUTHORS Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K. and  
 Tessier-Lavigne, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San  
 Francisco, CA 94143-0452, USA  
 FEATURES  
 source  
 1. 2697  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /tissue\_type="brain and ventral spinal cord"  
 /dev\_stage="18 day embryo and 13 day embryo"  
 1. 2697  
 /codon\_start=1  
 /product="transmembrane receptor UNC5H1"  
 /protein\_id="AAB57678.1"  
 /db\_xref="GI:2055392"  
 /translation="MAVRPGLWPVLGIIVLAALRGSGAQQSATVANPVPANPDLLP  
 HFLVEDPVIIVKPKVLLVCKAVPATQIFKCNQEWVQVDHVIERSDSSGLPTM  
 EVRINVSROOVKEVLEQVCCQVAVSSGTTKSKAYIRIAYLRNFEEQELAKEV  
 SLEQGLVLCRPPGPIPAEVENRNEVDLPDENVYITREHSLVVRQARLADAN  
 YTCVAKNIIVARRRSTAAVIVYVGGWSTWMSVCSACGRGWKRSCTNPAPLN  
 GGAFCEGQNVQKTACATLCPDGSWSWSKACGLDCTHRSRSCDPAKNGRECE  
 RGADLDNRCTSDLCITASCPEDEVALYIGLVAVAVCLFLLALLGLIYCRKEGIDS  
 DVADSLTSGFPVSIKPSKADNPHTLIQDLSITTTTGYGSLCSRODGPSPQL  
 SNGHLLSLPGSGRHTLHSSPTSEAEDEVSLSTQNYFRSLPRGTNSMAYFNFLGG  
 RLMIPTNGISLLIPDAIERGKIYEILFLHKPEDVRLPAGCTLLSPVSCGPGV  
 LUTRVLAMDHCGPSPDSWRUKKQCEGSEWEDVLHGESESHLYYCOLEAGC  
 VYFTQLGRFALVGEALSVAAKRLFLFAVACTSLRYNIRVYCLHDTALKEAV  
 QLEKQGLGLIQEPRVLHFQDSYHNLRLSIHDVPSLWSKLLYSQEIPEYHIWGT  
 COYLKCTFLERINASTDLACKVWVWVGEGQSPNFINFTKDRFALLALESEG  
 GYPALVGSAPFKIPFLIRQKIISLDPPCSRGRADWRTLAQKHLHLSHLSFASKPSPT  
 AMNLNWEARHPFNGLQGLAAVAGLQGPDAGLFTVSEAC"

Alignment Scores:  
 Pred. No.: 1,28e-279 Length: 2697  
 Score: 4545.50 Matches: 852  
 Percent Similarity: 96.78% Conservative: 19  
 Best Local Similarity: 94.67% Mismatches: 26  
 Query Match: 94.96% Indels: 3  
 DB: 10 Gaps: 3

US-09-970-944-2 (1-899) x RN087305 (1-2697)



QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeuValLeuAlaAlaTrpLeu 20  
Db 1 ATGGCCGTCGGCCCGCGCTGTGGCCAGTCTCTCTGGGCATAGTCTCTCGCGCGCTGCTT 60  
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
Db 61 CGTGGTTCGGGTGCCAGCAGAGTGCACGGTGGCCATCCAGTCCCGCTGCGCAACCC 120  
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTrileValLysAsnLysPro 60  
Db 121 GACCTGTCGCCACCTTCCTGGTAGAGCTGAGACGCTGACATTTGTCAAGAAACAAGCCG 180  
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnPhePheLysCysAsnGlyGlu 80  
Db 181 GTGTGTGTGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTCAAGTGCATATGGGAA 240  
QY 81 TrpValArgGlnValAspHisValLeuArgSerThrAspGlySerSerGlyGluPro 100  
Db 241 TGGTTCGCCAGGTTCGATCAGTAATTAACGCAGCACCCAGCAGCAGCGGATTGCCA 300  
QY 101 ThrMetGluValArgLysAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
Db 301 ACCATGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTGGGCTGGAG 360  
QY 121 GluTrpTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140  
Db 361 GAATACTGGTGGCCAGTGTGGCATGAGTCTCTGGGTACCACAAAGTCAAGAGGCC 420  
QY 141 TyrlleArgLysAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
Db 421 TACATCGGATGCTATTTCGCGAAGACTTTGAGCAGGAGCCACTGGCCAAAGAGTG 480  
QY 161 SerLeuGluGlnGlyLeuValLeuProCysArgProProGluGlyLeuProAlaGlu 180  
Db 481 TCACTGAGCAAGCATTTGTACTACTTGTGCCCCCGCCAGAGGAATCCCCCGAGCTGAG 540  
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTrile 200  
Db 541 GTGAGTGGCTTCGAAATAGAGACTCTGTGACCCCTCCCTCGATCCCAATGTGTACATC 600  
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTrpThr 220  
Db 601 ACGCGGAGACACGCTAGTCTGGGTGAGCCCGCTGGCCGACACGCGCAACTACACC 660  
QY 221 CysValAlaLysAsnLysValAlaArgArgArgSerAlaSerAlaAlaValLeuValTrp 240  
Db 661 TGTGTGCCAAGAACATCGTAGCCGTGCGCGAAGCACTCTGCGAGCGGTCAATTTAT 720  
QY 241 ValAsnGlyGlyTrpSerThrTrpThrTrpSerValCysSerAlaSerCysGlyArg 260  
Db 721 GTGAACGCTGGTGGTGCAGCTGAGTGGTCCGTCCTGAGCCCGCAGCTGTGGGCT 780  
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
Db 781 GGCTGGCAGAAACGAGCCGAGCTGACCAACCGGCACCTCTCAACGGGGCGCCCTTC 840  
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
Db 841 TGTGAGGCGCAGATGTC---CAGAAAACAGCCTGCGCCACTCTGTGCGCAGTGGATGG 897  
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
Db 898 AGCTGGAGTTCGTGGAGTAAGTGTGACGCTGTGGGCTGTGACTGCACCCACTGGCGGAGC 957  
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
Db 958 CGCGAGTGTCTGACCCAGCAGCCCGCAATGGAGTGTGAGGAGTGTGGGTCTGCACCTG 1017  
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
Db 1018 GACACCCGCACTGTACCACTGCTGCTGCGACACCGCTCTTGTGCCCGGAGACGTG 1077  
QY 361 AlaLeuTyrrValGlyLeuLeuAlaValAlaValCysLeuValLeuLeuLeuValLeu 380

Db 1078 GCTCTCTACATCGCCTTGTGGCTGTGGCTGTGGCTCTCTCTGTGTGTGTGGCCCTT 1137  
QY 381 IleLeuValTyrrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400  
Db 1138 GGACTCATTTTACTGTGCAAGAGAGAGGCTGGACTCCGATGTGGCCGACTCGTCAATC 1197  
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
Db 1198 CTCACTCTGGGCTTTCAGCTGTGAGCATCAAGCCAGCAAGAGCAGACACCCCACTG 1257  
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTrpThrGlnGlySerLeuCysPro 439  
Db 1258 CTCACCATCCAGCAGACCTTCAGCACCCACCTACCACTTACCAGGCGAGTCTATGTTCG 1317  
QY 440 ArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
Db 1318 AGCAGGATGGACCCAGCCCAAGTTCCAGCTCTCTAATGTCACCTGCTCAGCCCACTG 1377  
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479  
Db 1378 GGGAGTGGCCGCTATAGTGTGACACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437  
QY 480 SerArgLeuSerThrGlnAsnTyrrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
Db 1438 TCCGCGCTCTCCACCCCAAACTACTTTCGTTCCCTGCGCCGCGCACCCAGCAACATGGCC 1497  
QY 500 TyrrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
Db 1498 TACGGGACCTTCAACTTCCTCGGGGCGGCTGATGATCCCTAATACGGGATCAGGCTC 1557  
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrrGluIleTyrrLeuThrLeuHis 539  
Db 1558 CTATATCCCGGATGCCATCCCGAGAGAAAGATCTACGAGATCTACCTCACACTGCAC 1617  
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
Db 1618 AAGCAGCAAGACGTGAGTGTGCCCTAGTGGCTGTGAGACCCCTGCTGAGTCCAGTGGTT 1677  
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
Db 1678 AGCTGTGGGCGCCAGAGTCTCTGCTACCCGCGCAGTATCTCTGATGATGAGACACTGT 1737  
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599  
Db 1738 GGAGAGCCAGCCCTGACAGCTGGAGTCTGGCTCAAAAGACAGTCTCTGCGAGGCGAGT 1797  
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrrTyrrCysGln 619  
Db 1798 TGGGAG--GATGTGCTGCACCTTGGTGAGGAGTCACTTCCACCTCTACTACTGCCAG 1854  
QY 620 LeuGluAlaSerAlaCysTyrrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
Db 1855 CTGAGGCGCGGGCTGCTATGTCTTCAGGAGCAGCTGGGCGCGCTTTCCTCTGGTAGGA 1914  
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659  
Db 1915 GAGGCCCTCAGCTGGCTGCCACCAAGCGCTCAGGCTCTCTGTTTGTCTCCGCTGCC 1974  
QY 660 CysThrSerLeuGluTyrrAsnIleArgValTyrrCysLeuHisAspThrHisAspAlaLeu 679  
Db 1975 TGTAGTCCCTTGTAGTACAACATCCGAGTCTGCTTACACGACCCAGCAGCAGCTCTC 2034  
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgVal 699  
Db 2035 AAGGAGGTGGTGGCTGGAGAGCAGTGTGTGAGCAGCTGATCCAGGAGCTCGCTC 2094  
QY 700 LeuHisPheLysAspSerTyrrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
Db 2095 CTGCACTTCAAAGACAGTTACCAACCTACGCTCTCTCCATCCAGCAGTGGCCAGCTCC 2154  
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrrGlnGluIleProPheTyrrHisIleTrpAsn 739



```

Db      2155 CTGTGGGAAGACAAAGCTACTTGTACAGTACCAGGAGATCCCTTTTACCAATCTGGAAAC 2214
Qy      740 GlyThrArgGlyLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
      |||.....|
Db      2215 GGCACCCAGCAGTATCTGCATCTGACCTTACCTTGGAGCGCATCAACGCCAGCACCAGC 2274
Qy      760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerile 779
      |||.....|
Db      2275 GACCTGGCTGCAAGGTGTGGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATC 2334
Qy      780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
      |||.....|
Db      2335 AACITCAACATCCTAAGGACACAAAGGTTTGTGTAATTGTTGGCTCTGGAGAGTCAAGGG 2394
Qy      800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
      |||.....|
Db      2395 GGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATTCCTCTCATTCGGGCAAAAG 2454
Qy      820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
      |||.....|
Db      2455 ATCATCGCCAGTCTGGAGCCCAACCTCGCAGCGGGGGCCGACTGGAGAACTCTAGCCAG 2514
Qy      840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
      |||.....|
Db      2515 AAACCTTCACCTGGACAGCCATCTTAGCTTCTTTCCTCCAAAGCCCGACCCCTACAGCCATG 2574
Qy      860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
      |||.....|
Db      2575 ATCTCAACCTATGGAGGACGCGACCTTCCCAACCGCAACCTCGGCCAGCTGGCAGCA 2634
Qy      880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
      |||.....|
Db      2635 GCTGTGCGCGACTGGGCCAACACAGATGCTGGCTCTTTCAGGTGTGGAGGCCGAGTGT 2694

RESULT 7
BD057524
LOCUS      BD057524              3014 bp    DNA    linear    PAT 27-AUG-2002
DEFINITION Netrin receptors.
ACCESSION   BD057524
VERSION     BD057524.1 GI:22603130
KEYWORDS    JP 2001505062-A/1.
SOURCE      synthetic construct
            synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 3014)
            Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
AUTHORS     Netrin receptors
TITLE       Patent: JP 2001505062-A 1 17-APR-2001;
            THE REGENTS OF THE UNIV OF CALIFORNIA
JOURNAL     EN JP 2001505062-A/1 17-APR-2001
COMMENT     PD JP 2001505062-A/1
            PF 19-FEB-1998 JP 1998536840
            PR 19-FEB-1997 US 08/808982
            PI MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
            PI MASU,
            PI KAZUKO KEINO MASU
            PC C07K14/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
            Strandedness: Double;
            CC Topology: Linear;
            PH Key Location/Qualifiers.

FEATURES             Location/Qualifiers
     source           1..3014
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"

ORIGIN
Alignment Scores:
Pred. No.:          1,48e-279      Length:          3014
Score:              4545.50        Matches:          852
Percent Similarity: 96.78%         Conservative:
Best Local Similarity: 94.67%         Mismatches:        26
Query Match:        94.96%          Indels:           3

```

```

DB:      6              Gaps:      3
US-09-970-944-2 (1-899) x BD057524 (1-3014)
Qy      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
      |||.....|
Db      1 ATGGCCGTCCGCCCGCGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCCCTTGGCTT 60
Qy      21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
      |||.....|
Db      61 CGTGGTTCGGGTGGCCAGCAGTGCACGTGGCCAAATCCAGTGCCTGGTGGTGGTGGTGGT 120
Qy      41 AspLeuLeuProHisPheLeuValGluProGlnAspValTyrIleValLysAsnLysPro 60
      |||.....|
Db      121 GACCTGTGCTGCCCTCTCTCTGGTAGAGCTGAGGACGTGTACATTGTCAAGTGCATGGGAA 180
Qy      61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
      |||.....|
Db      181 GTGTGTGTGTGTGTCAGAGGTGTGCTTGCACCCAGATCTTCTTCAAGTGCATGGGAA 240
Qy      81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
      |||.....|
Db      241 TGGGTCCGCCAGGTGTGATCAGTAAATTGAACGACGACCGACAGCAGCAGCGGATGGCA 300
Qy      101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
      |||.....|
Db      301 ACCATGAGGTGCGGTATCAACGTATCGAGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360
Qy      121 GluTrpTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
      |||.....|
Db      361 GAATACTGGTGGCCAGTGTGTGGCATGTGCTCGGTACCCACCAAAAGTCAAGAGGCC 420
Qy      141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
      |||.....|
Db      421 TACATCCGATTCCTATTTTCGCAAGAACTTTGAGCAGCAGGACCTTGGCCAGGAAGTG 480
Qy      161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
      |||.....|
Db      481 TCACTGGAGCAAGCAATTGTACTACCTTGTGCGCCCCCAGAGGAATTCCTCCAGTGTGATC 540
Qy      181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
      |||.....|
Db      541 GTGAGTGGCTTCGAAATGAGGACCTCTGTGGACCCCTCCCTCGATCCCAATGTGTATC 600
Qy      201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220
      |||.....|
Db      601 ACCTGGGAGCACAGCCTAGTCTGTGCGTCAAGCCCGCTGGCCGACACGCGCAACTACACC 660
Qy      221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240
      |||.....|
Db      661 TGTGTGGCCAAAGAACATCGTAGCCCGTCCGCGAAGCACCTCTGACGCGTCAATTGTTAT 720
Qy      241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260
      |||.....|
Db      721 GTGAACGTTGGTGGTGTGCGAGTGTGAGTGTCTCGTCTGACGCGCCAGCTGTGGCGGT 780
Qy      261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
      |||.....|
Db      781 GGCTGGCAGAAACGAGCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGCGGCTTC 840
Qy      281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
      |||.....|
Db      841 TGTGAGGGGCAAGTGTCTCAGAAACAGCGCTGCGCCACTCTGTGCCAGTGTGGATGGG 897
Qy      301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
      |||.....|
Db      898 AGCTGGAGTTCGTGGAGTAAAGTGTGTCAGCTGTGGCTTTGACTGCACCCACCTGGCGGAGC 957
Qy      321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
      |||.....|
Db      958 CGCGAGTGTCTGTGACCCAGCAGCAGCCCGCAATGGAGGTGAGAGTGTGGGGGTGCTGACCTG 1017
Qy      341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360

```

```
Db 1018 GACACCGCAACGTGACCAAGTACCTCTGCTGCACACCGCTCTCTGCCCCGAGAGCGTG 1077
Qy 361 AlaLeuTyrValGlyLeuLeuAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
Db 1078 GCTCTCTACATCGCCCTTGCTGCTGCTGCTGCTGCTCTCTCTCTGCTGCTGCTGCT 1137
Qy 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400
Db 1138 GGACTCATTTACTGTCGAAAGGAAGGAGGCTGAGCTCCGATGCTGGCGGAGCTGCTCATC 1197
Qy 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1198 CTCACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCGAGCAAGAGCAACCCCCACCTG 1257
Qy 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThr 439
Db 1258 CTCACCATCCAGCCAGACCTCAGCACCCACCTACACCTACAGGGGAGTCTATGTCG 1317
Qy 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1318 AGCAGGATGACCCAGCCCAAGTTCAGCTCTCTAATGCTCACTGCTCAGCCCACTG 1377
Qy 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479
Db 1378 GGGAGTGGCCGCCATACGTTGCACACAGCTTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437
Qy 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1438 TCCGCGCTCTCCACCCAAACTACTTTCGTTCCCTGCGCCGCGCACCAACATGGCC 1497
Qy 500 TyrGlyThrPheAsnLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1498 TACGGGACCTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557
Qy 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db 1558 CTCATACCCCGGATGCCATCCCCGGAGGAAGATCTACGAGATCTACCTCAGACTGCAC 1617
Qy 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1618 AAGCCAGAGACGTGAGGTTCGCCCTAGCTGGCTGTGACAGCCCTGCTGAGTCCAGTCTT 1677
Qy 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1678 AGCTGTGGCCCCCAGGAGTCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGT 1737
Qy 580 GlyGluProSerProAspSerTyrSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1738 GGAGAGCCCGCCCTGACAGCTGGAGTCTGGCCCTCAAAAGCAGTCTCGGAGGGCAGT 1797
Qy 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619
Db 1798 TGGGAG---GATGTGCTGCACCTTGGTGAGGAGTCACTTCCCACTCTACTACTGCGCAG 1854
Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGGCGGGCGCTGTATGCTTCACGGAGCAGCTGGGGCGGCTTTGCCCTGTGTAGGA 1914
Qy 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db 1915 GAGGCGCTCAGGTGCTGCCACCAAGCGCTCAGGCTCTCTGTTGCTCCCGTGCGC 1974
Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTACGTCCTTGTAGTACAACATCCGAGTACTGCTACAGCAGCACCCACGACGCTCTC 2034
Qy 680 LysGluValValGlnLeuLysGlnLeuGlyGlyGlnLeuLeuGlnGluProArgVal 699
Db 2035 AAGGAGGTGGTGCAGCTGGAGAGGAGCTAGGTGGACAGCTATCCAGAGGCTTCGCTC 2094
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
```

```
2095 CTGCACCTTCAAGACAGAGTGTACCAACAACCTACGTCTCTCCATCCAGCAGCTGCCAGCTCC 2154
Qy 720 LeuTyrLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAAGAGCAAGTACTTGTCACTACAGAGATCCCTTTTACCAATCTGGAAC 2214
Qy 740 GlyThrGluArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGCAGTATCTGCACCTTCACTCCCTGGAGGCGCATCAAGCCAGCAGCCAGC 2274
Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGAGATGGCAGAGCTTCAACATC 2334
Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2335 AACCTCAACATCATCAAGGACCAAGGTTTGCTGAATTGTGGCTCTGGAGAGTGAAGGG 2394
Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCCCTCATTGGGCAAAAG 2454
Qy 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATCATCGCAGTCTGGACCCACCTGTCAGCGGGCGGCGGCTGGAGAACTCTAGCCGAG 2514
Qy 840 LysLeuHisLysAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2515 AAACCTTCACTTGGACAGCCATCTTAGCTTCTTTCCTTCCAAAGCCAGCCCTAGAGCCATG 2574
Qy 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db 2575 ATCTCTCAACCTATGGGAGGACAGGCACTTCCCCAACGCGCAACCTCGGCCAGTGGCAGCA 2634
Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2635 GCTGTGGCGGAGTGGGCCAACAGATGCTGGCCTCTTTCAGGTTGCGGAGGCGGAGTGT 2694

RESULT 8
AX367094 3580 bp DNA linear PAT 16-FEB-2002
LOCUS Sequence 13 from Patent WO0198354.
ACCESSION AX367094
VERSION AX367094.1 GI:18855296
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B., Tang,X.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R., Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O., Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and Sanjanwala,M.S.
TITLE Receptors
JOURNAL Patent: WO 0198354-A 13 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 3580
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 6052371CB1"

Alignment Scores:
Pred. No.: 6,64E-269 Length: 3580
Score: 4379.50 Matches: 838
Percent Similarity: 93.22% Conservative: 1
Best Local Similarity: 93.11% Mismatches: 2
Query Match: 91.49% Indels: 59
DB: Gaps: 3
```

US-09-970-944-2 (1-899) x AX367094 (1-3580)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20  
Db 4 ATGGCGTCCCGCGCGCTGTGGCAGCGTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 63

QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAlaAsnProValProGlyAlaAsnPro 40  
Db 64 CGCGCTCGGTGCCAGCAGGTGCCACCGTGGCCAAACCCAGTGGCTGGTGGCCAAACCCG 123

QY 41 AspLeuLeuProHisPheLeuValGlnProGluAspValTyrIleValIleValIleValIleVal 60  
Db 124 GACCTGCTTCCCACTTCTCTGGTGGAGCCCGAGGATGTATCATCTCAAGAAACAAGCCA 183

QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
Db 184 GTGCTGCTGTGTCAAGGCGCTGCCCGCCAGGAGATCTTCTTCAAGTGAACGGGGAG 243

QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
Db 244 TGGGTGCCAGTGGACCATGATCGAGGCGACACAGACGGAGGAGTGGGCTGCC 303

QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
Db 304 ACCATGGAGGTCCGATTAATGTCTCAAGGCAGCAGGTTCGAGAAGTGTTCGGGCTGGAG 363

QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140  
Db 364 GAATACTGGTGCCAGTGGTGGCATGAGGCTCTCGGGGCACCCACCAAGAGTCAAGAAGGCC 423

QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
Db 424 TACATCCGATAGCTATTGGCAAGAACTTCAGCAGGAGCGCGCTGGCCAGGAGGTG 483

QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGlu 180  
Db 484 TCCCTGGAGCAGGCGATCGTGTCCCTCGCTCCACCGAGGGCATCCCTCCAGCGCGAG 543

QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
Db 544 GTGAGTGGCTCGGAACAGGAGGACTGTGGACCCGCTCCCTGGACCCCAATGTATATC 603

QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
Db 604 ACCGGGAGCACAGCTGTGTGGCGACAGCGCGCGCTTCTGACCGCCAACTACACC 663

QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
Db 664 TGGGTGGCCCAAGAACATCGTGGCACGTCGCGCAGCGCCCTCCGCTGTGTCTATCTATC 723

QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
Db 723 ----- 723

QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
Db 723 ----- 723

QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
Db 724 -----GTGGACGGC 732

QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
Db 733 AGCTGAGCCCGTGGAGCAAGTGGTGGCGCTGTGGCTGGACTGGACCCCACTGGCGGAGC 792

QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
Db 793 CGTGAGTGTCTGACCCAGGACCCCGCAACGGAGGGAGGAGTGGCCAGGGCACTGACCTG 852

QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
Db -----

853 GACACCCGCAACTGTACCAGTGCCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG 912

361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380

913 GCCCTATGTGGGCTCATCGCGTGGCGCTCTGCTGTGGTCTGCTGCTGCTGCTGCTC 972

381 IleLeuValTyrCysArgGlyLysGlyLeuAspSerAspValAlaAspSerSerIle 400

973 ATCTCTGTTTATTGCCGGAAGAGGGGCTGGACTAGATGTGGCTGACTGCTGCTCAAT 1032

401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420

1033 CTCACCTCAGGCTTCCAGCGCTCAGCATCAAGGCCAGCAAGAACACACCCCACTG 1092

421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr 439

1093 CTCACCATCCAGCGGACCTCAGCACCCACACCACTACCAGGGCAGTCTCTGTCTCC 1152

440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459

1153 CGGAGGATGGGCGCCAGCCCAAGTTCAGCTCACCAATGGGCACCTGCTCAGCCCTG 1212

460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal 479

1213 GGTGGCGCGCCACACACTGCACACAGCTCTCCCACTCTGAGGCGGAGGAGTTCGTC 1272

480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499

1273 TCCCGCTCTCCACCCAGAACTACTTCGCTCCCTGCTCCCGAGGACACCAATGAC 1332

500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519

1333 TATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACAGGAATCAGGCTC 1392

520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539

1393 CTCATCCCCCAGATGCCATACCCCGAGGAGAAATCTATGAGATCTACCTACGCTGCAC 1452

540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559

1453 AAGCGGAAGACGTGAGTTGCCCTAGCTGGCTGTTCAGACCCCTGCTGAGTCCCATGTT 1512

560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579

1513 AGTGTGGACCCCTCGCTGCTGCTCACCAGGCGAGTCACTCTGGCTATGGACCACTGT 1572

580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGlyGlySer 599

1573 GGGGAGCCAGCCCTGACAGCTGGAGCTCGGCTCAAAAAGCAGTCTGTCGAGGGGAGC 1632

600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGln 619

1633 TGGGAG---GATGTGTGCACCTGGCGAGGCGCGCTCCCACTCTACTACTGCGAG 1689

620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639

1690 CTGGAGCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1749

640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659

1750 GAGGCCCTCAGCGTGGCTGCCCAAGCGCTCAAGCTGCTTCTGTTTGGCCCGGTGGCC 1809

660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679

1810 TGCACCTCCTCGAGTACAACATCCGGGTCTACTGCTCTGATGACACACCCACGATGACATC 1869

680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699

1870 AAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCCAGGAGGCACCGGTC 1929

700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719

1930 CTGCACCTTCAAGACAGTTTACCACACCTTCCGCTTATCCATCCAGATGTGCCCGCTCC 1989



/note="ZUS; Region: Domain present in ZO-1 and Unc5-like netrin receptors"  
 /db\_xref="CDD:smart00218"  
 2495.. 2684  
 misc\_feature

/note="DEATH; Region: DEATH domain, found in proteins involved in cell death (apoptosis). Alpha-helical domain present in a variety of proteins with apoptotic functions. Some (but not all) of these domains form homotypic and heterotypic dimers"  
 /db\_xref="CDD:smart00005"

## ORIGIN

## Alignment Scores:

Pred. No.: 1-93e-262 Length: 3844  
 Score: 4278.50 Matches: 809  
 Percent Similarity: 92.00% Conservative: 19  
 Best Local Similarity: 89.89% Mismatches: 13  
 Query Match: 89.38% Indels: 59  
 DB: 10 Gaps: 3

US-09-970-944-2 (1-899) x BC058084 (1-3844)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20  
 DB 252 ARGCGCGTCCGCGCCGCGCTGTGGCAGCGTCTCTGGCAGTAGTCTCACTGCGCTGGCTT 311  
 QY 21 ArgGlySerGlyAlaGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
 DB 312 CGTGGTTCGGGTGCCACAGAGTCCACAGTGGCCAAACCAGTGCCCTGGTGGCCAAACCG 371  
 QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60  
 DB 372 GACCTGTCTGCCCACTTCTCGTAGCGCGGAGGACGTGTACATGTCAAGAACAGAGCCC 431  
 QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
 DB 432 GTGCTGTGGTGTGCAAGGCTGTGCCGCCACCCAGATCTTCTCAAGTGCAAGCGGAA 491  
 QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
 DB 492 TGGGTTGCCAGAGTCGATCAGTCATTGAACGCGAGCATGACGCGCAGCGAGATTCGCCA 551  
 QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
 DB 552 ACCATGAGGTCGGATCAACGTATCAAGGACGAGTCCGAGAAGTGTTTGGGCTGGAG 611  
 QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140  
 DB 612 GAGTACTGGTGCCAGTGTGTGGCATGGAGCTCTCTCAGGAACACCAAGAGCCAGAGGCC 671  
 QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
 DB 672 TACATCCGGATTGCTATTTCGCAAGAACCTTTGAGCAGGAGCGCTGGCCCAAGGAAGTG 731  
 QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180  
 DB 732 TCACCTGAGCAAGGCAATTGTGTACTTGTGCCCGCCGAGAGAAATCCCGCCAGCTGAG 791  
 QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
 DB 792 GTGGAGTGGCTCCGAAATGAGGACCTCGTGAGACCCCTCCCTCGACCCCAATGTGTATC 851  
 QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
 DB 852 ACAGGAGACAGCTATGTCTGTGGCAGCGCCGCGCTCGACCCCACTCCCACTACACC 911  
 QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr 240  
 DB 912 TGGCTGGCCAAAGAACATCGTGGCCCGTCCGGAAGCGCTCTCGGCGCTCATTTGTTAT 971  
 QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
 DB 971 ----- 971

QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
 DB 971 ----- 971  
 QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
 DB 972 -----GTGGATGGG 980  
 QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
 DB 981 AGCTGGAGCCCATGGAGTAAGTGTACGCTCGCGGCTTGACTGCACCCACTGGCGAGC 1040  
 QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
 DB 1041 CGGAGTGTCTCCGACCCAGCGCCCAACGAGGTGAGGAGTGC CGGGGTCTGACCTG 1100  
 QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
 DB 1101 GACACCCGCAACTGTACCAAGTACCTCTGCGTGCACACCTCTTCGCGCCCGGAGACGTG 1160  
 QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
 DB 1161 GCTCTCTACATCGGCTCGTCCGCGTGGCGGTGGCTTCATCTTCTGCTGCTGCTGCTC 1220  
 QY 381 IleLeuValTyrCysArgLysGlyGlyLeuAspSerAspValAlaAspSerSerIle 400  
 DB 1221 GTCCTCATCTACTGCGCAAGAGAGAGACTGGACTCAGACGTGGCTGACTCATCCATC 1280  
 QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
 DB 1281 CTTTACCTCAGGCTTCCAGCCTGTGAGCATCAAGCCCAAGCAAGCAGACAATCCCCATCTG 1340  
 QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439  
 DB 1341 CTCACCATCCAAACCGGACCTCAGCACCAACAGCAGCCACTACAGGGGAGGCTGTGTGCC 1400  
 QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
 DB 1401 CGGAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGTGTACCTGTCTCAGCCCATG 1460  
 QY 460 GlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGluGluPheVal 479  
 DB 1461 GGCAGTGGCGGCCATACGCTGCACACAGCTCCCGCCACCTCTGAGGCTGAGGACTTCGTC 1520  
 QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
 DB 1521 TCCGCGCTCTCCACCCCAAACTACTTTCTGTTCTCTGCCCCGCGGTACAGCAACATGCC 1580  
 QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
 DB 1581 TATGGAGCTTCAACTTCTCGGGGCGGCTGATGATCCCTAACACAGGAATCAGCCTC 1640  
 QY 520 LeuIleProProAspAlaIleProArgGlyIleTyrGluIleTyrLeuThrLeuHis 539  
 DB 1641 CTATACCCCGGAGCGCCATCCCGCGAGGAAGATCTACAGAGATCTACCTCACTCTGCAC 1700  
 QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
 DB 1701 AAGCCAGAACGCTGAGTTGCCCTAGCTGGCTGTGACACCTCTGAGTCTTATCGTT 1760  
 QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
 DB 1761 AGCTGTGGGCCCCAGAGTCTGTCTCACCGGCCAGTCACTCTTGCATGGACCACTGC 1820  
 QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysGlnSerCysGluGlySer 599  
 DB 1821 GGGAGCCAGTCCCGACAGCTGAGCTCGGCTCAAAAAGCAGTCTCTGTAGGGGAGC 1880  
 QY 600 TrpGluAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619  
 DB 1881 TGGGAG---GACGTGTGACCTTGTGAGGAGTGGCCCTCTCATCTACTACTGCCAG 1937

QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
 Db 1938 CTGAGGCGGGGCTGCTATGCTTTACAGAGCAGTAGGCGCTTTGGCCCTGGTGGGA 1997  
 QY 640 GluAlaLeuSerValAlaAlaAlaValArgLeuLeuLeuPheAlaProValAla 659  
 Db 1998 GAGGCCCTCAGCTGGCTGCCACCAAGCGCTCAGGCTCTCTGTTGGCCCTGTGGCC 2057  
 QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
 Db 2058 TGTACGTCCTCGAGTACCAACATCCGAGTGTACTGCTGCAGCAGCACCCAGATGCTCTC 2117  
 QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuGlnGlnProArgVal 699  
 Db 2118 AAGGAGGTGTGTGAGTGGAGAAGCAGCTGGGTGGACAGCTGTATCCAGGAGCCCGCTGTC 2177  
 QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
 Db 2178 CTGCACCTCAAGACAGTATACCAACCTAGCTGTCTGCATCCAGCAGCTGCCAGCTCC 2237  
 QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
 Db 2238 CTGTGGAAGAGCAAGCTCTGTGCTAGCTACCAAGAGATCCCTTTTACCATCTGGAAT 2297  
 QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
 Db 2298 GGCATCTACAGATATCTGCTAGCTACCAAGAGATCCCTTTTACCATCTGGAAT 2357  
 QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779  
 Db 2358 GACCTGGCTGCAGAGTGTGGTGTGGCAGGTGGAGGAGATGGACAGAGCTTCAACATC 2417  
 QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
 Db 2418 AACTTTAATCACTACCAAGACAGAGTGTGTGAAATGCTGGCTCTGGAGAGTGAAGG 2477  
 QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
 Db 2478 GGGGTCCAGGCGCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAAG 2537  
 QY 820 IleLeuSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839  
 Db 2538 ATCATACAGCTGACCCACCTTCAGCGCGGCGCGAGTGGCGAATCTAGCCAG 2597  
 QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859  
 Db 2598 AAACCTTCACTGAGCAGCATCTTAGCTTCTTTGGCTCCCAAGCCAGCCCTACAGCATG 2657  
 QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaLa 879  
 Db 2658 ATCTCTCAACCTATGGAGCGGGCAGCTTCCCAAGCGCACTCGCCAGCTGGCCGA 2717  
 QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
 Db 2718 GCTGTGGCGGACTGGCCAGCAGCATGTGGCTCTTCAAGCTGTGAGAGCGGAGTGC 2777  
 RESULT 10  
 AK122575  
 LOCUS Mus musculus mRNA 4294 bp mRNA linear ROD 15-MAR-2003  
 DEFINITION Mus musculus mRNA for mKIAA1976 protein.  
 ACCESSION AK122575  
 VERSION AK122575.1 GI:28972881  
 KEYWORDS FLI-CDNA.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,  
 Nakajima,D., Nagase,T., Ohara,O. and Koga,H.  
 Prediction of the coding sequences of mouse homologues of KIAA  
 gene: II. The complete nucleotide sequences of 400 mouse  
 KIAA-homologous cDNAs identified by screening of terminal sequences

of cDNA clones randomly sampled from size-fractionated libraries  
 DNA Res. 10, 35-48 (2003)  
 2 (bases 1 to 4294)  
 Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.  
 Direct Submission  
 Submitted (07-FEB-2003) Hisaaki Koga, Kazusa DNA Research  
 Institute, Laboratory for Genome Informatics; 2-6-7  
 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan  
 (E-mail:mouse@kazusa.or.jp, Tel.:81-438-52-3919, Fax:81-438-52-3918)  
 The CREATE program supported by Japan science and technology  
 corporation; cDNA full insert sequencing: Kazusa DNA Research  
 Institute; cDNA library construction, clone selection and 5'- &  
 3'-end one pass sequencing.  
 FEATURES  
 Location/Qualifiers  
 1..4294  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="mbh03993"  
 /tissue\_type="brain"  
 /dev\_stage="adult"  
 /note="vector:modified pBC SK+"  
 1..4294  
 /gene="mKIAA1976"  
 /cDS\_start=1  
 /evidence="not experimental"  
 /product="mKIAA1976 protein"  
 /protein\_id="BAC65857.1"  
 /db\_xref="GI:28972882"  
 /translation="SLVTPALLKTTLSAASSGPEDVALYGLVAVAVCLILLLVV  
 LIYCRKKEGLSDVADSSILTSFGQPSVSKSKADNHLIIQDPLSTTTTQGSGLC  
 PRQGPSPFQLEKQGLQIEPRVHLFKDSYHNLRISIHDSVSSKMKLLYSQOE  
 IPFYHNGTQYLLHCTFLERVNASDLSKVMWQVEGDSFNFINFKIDTRF  
 AEMALISEGVPALVGPFAFKIPFLIRKIIITSLDPPCSRGAOWRTLAQKLDHSL  
 SFFASKPSPTAMILNLWEARHFPNGNLQQLAAAVAGLQPDAGLFTVSEAC"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,216-182 Length: 4294  
 Score: 3024.50 Matches: 632  
 Percent Similarity: 49.09% Conservative: 15  
 Best Local Similarity: 47.95% Mismatches: 15  
 Query Match: 63.18% Indels: 657  
 DB: 10 Gaps: 4  
 US-09-970-944-2 (1-899) x AK122575 (1-4294)  
 QY 2 AlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArg 21  
 Db 2 GCGCTCCGCGCGCGCTGTGGCCAGCGCTCTGGGCGATAGTCTCTACTGCGCTGGCTTGT 61  
 QY 22 GlySerGlyAlaGlnSerAlaThrValAlaAlaAsnProValProGlyAlaAsnProAsp 41  
 Db 62 GGTTCGGGTGCCAGCAGAGTGCACAGTGCACACCCAGTGCCTGGTCCCAACCGGAC 121  
 QY 42 LeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysProVal 61  
 Db 122 CTGCTGCCCACTCTCTGGTAGCGGAGGAGCTGTATCTGTCAAGAACAGCGCGGTG 181  
 QY 62 LeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrp 81  
 Db 182 CTGCTGGTGTGAAGCTGTGCCGCCACCCAGATCTCTTCAAGTGCACCGGGATGG 241  
 QY 82 ValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluProThr 101  
 Db 242 GTTCCAGGTGCATCAGTCTTGAACGACAGTGCACGAGCAGCAGCAGGATGCCAAC 301  
 QY 102 MetGluValArgIleAsnValSerArgGlnValGluLysValPheGlyLeuGluGlu 121

Db 302 ATGGAGGTCGGATCAACGTTATCAAGGCAGCAGGTCTGAGAAAGTCTTTGGGCTGGAGGAG 361  
 QY 122 TyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyr 141  
 Db 362 TACTGGTGCAGTGTGTGGCATGGAGCTCTCTAGGAACACCAAAAGCCAGAAGCCCTAC 421  
 QY 142 IleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSer 161  
 Db 422 ATCCGATTGCTATTTCGCAAGAACTTTGAGCAGGAGCGCTGGCCCAAGAAAGTGCA 481  
 QY 162 LeuGluGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGluVal 181  
 Db 482 CTGAGCAAGGCATGTGTCTACCTTGTGCGCCCGCCGGAAGAAATCCCGCCAGCTGAGGTG 541  
 QY 182 GluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThr 201  
 Db 542 GAGTGGCTCCGAATAGAGCACTGTGGACCCCTCCCTCGACCCCAATGTGTATCATACA 601  
 QY 202 ArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCys 221  
 Db 602 CGGAGCAGACAGCTAGTGTGCGGCAGCGCGCGCTGGCCGACACTGCCAACTACACCTGC 661  
 QY 222 ValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrVal 241  
 Db 662 GTGGCCCAAGAAACATCGTGGCCCGCTGCGGAAGCGCTCTGCGGCGCTCATTTGTTATGTG 721  
 QY 242 AsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGly 261  
 Db 722 AACGGTGGGTGCAGTGTGACCGAGTGTGCTGTCAGCGCCAGCTGTGGCGGTGGC 781  
 QY 262 TrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCys 281  
 Db 782 TGGCAGAAACGGAGCGAGCTGCACCAACCCCGCACCTCTCAACGGGGCGCTTCTGT 841  
 QY 282 GluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySer 301  
 Db 842 GAGGGGCAGAAATGTC---CAGAAACACCGCTGTCGCCAGTGGATGGGAGC 898  
 QY 302 TrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArg 321  
 Db 899 TGGAGCCATGGAGTAGTGTGAGCTGCGGGCTTGACTGCACCCACTGGCGGAGCGG 958  
 QY 322 GluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAsp 341  
 Db 959 GAGTGTCTCCGACCCAGCGCCCGCAACGGAGGTGAGGAGTGGCGGGTGTGACTGGAC 1018  
 QY 342 ThrArgAsnCysThrSerAspLeuCysValHisSerAla----- 354  
 Db 1019 ACCCGCAACTGTACCACTGACGTGACCTTGCTGTGCACAGTGA-GTCTTTCTCTGCCCCAGGGGCTC 1077  
 QY 354 ----- 354  
 Db 1078 ACCTCGCTGTAGATGCCCTATGGCTGTGGTTCAGGCGAGCCACTCTCTCGCTGCCTGTGTC 1137  
 QY 354 ----- 354  
 Db 1138 TCCAAAGCTTGGAGCCAGGGGTCCCTTTCATCTCTAGCACACTCTCTCCCTGTACACA 1197  
 QY 354 ----- 354  
 Db 1198 GACACACCACCTTCCCGCTTGAGACACACAGATGTGAATCTTTGACCCCTGTATACCAA 1257  
 QY 354 ----- 354  
 Db 1258 CTGCTCCCCCGGGCTTTGAAATTATCCCGTCCCGCAAGCTCACATAAGTCTCTGTGGG 1317  
 QY 354 ----- 354  
 Db 1318 GGTGCTGGGCTTGTGAGGCCCTGTTAATTACAGGCCATAACTAACAAATTAAGATCTCTG 1377  
 QY 354 ----- 354  
 Db 1378 TTGGATTTTTTTTTTCTTTAGACTACACCTCCCTCCCTGTTCTTCAATTAAAAACCTAA 1437

QY 354 ----- 354  
 Db 1438 TTATCTGTGTCAATTCCTTCTGTATGGACATTTCTTCCATATGCAAAACTCTTAAAAAT 1497  
 QY 354 ----- 354  
 Db 1498 CGATAGTGTCTTCTCCAAAGCAGTTGGTCCCCCACCACCCAGTTTCAGGGCTGGTGA 1557  
 QY 354 ----- 354  
 Db 1558 TACAAAGGGCTCACACTGTCTCTGGAACACCCCGAAGAACTGACCCCATGTATCTGCC 1617  
 QY 354 ----- 354  
 Db 1618 TATTTTCTATTCTATAATGAAGTAGCAAAAGAAATTTATTAGTTTATAGAGCTCAAAAT 1677  
 QY 354 ----- 354  
 Db 1678 CCAATAGCGTGTGCTGGCTCTGGCAAAATGTGTCTTCCACCCCAAGCCACATTTGCTTGG 1737  
 QY 354 ----- 354  
 Db 1738 TGGCAGATAGTACCACATCAGTACTGTAGACAAGAGAGAGATCACGTCTCAGGAAGAGAG 1797  
 QY 354 ----- 354  
 Db 1798 AAAGACAGACAGAGACAGGTAGACACAGGAAGATGGACAGCCCCCATGGTCTTAAGACA 1857  
 QY 354 ----- 354  
 Db 1858 GCACATAATGATGTAAATGCCTCCCCCATGGGGGCCACCTCTTAAAGGCGCCACTGCT 1917  
 QY 354 ----- 354  
 Db 1918 GCTCATATTGCCAATGGGACACACACCTGTCTCAATATGAATCCTTAGGAGGATGTGC 1977  
 QY 354 ----- 354  
 Db 1978 AAACTATAGCACCATACTTAGGGTCTCTCAGGAGTTTCACTTGTCCCCAGACACCCCTGACC 2037  
 QY 354 ----- 354  
 Db 2038 CCCAGCCTAGAGCCAAAGGATATACCACTCACCCCGGTGAAGTAGTCTTCCCTGCTGC 2097  
 QY 354 ----- 354  
 Db 2098 TCCACTCTGCTGTGAGGGTGTGACTTGTCTCAATCCCAAGCAAGGTGTGTGGGTATAG 2157  
 QY 354 ----- 354  
 Db 2158 AACTAGGATGTACCCACCAATACCAAGCTCCCTGCGCACGTGCCACTGTGCCAGGGGTGC 2217  
 QY 354 ----- 354  
 Db 2218 TGAGATCTATTCTCTGGGGATAAACCATATGGGCTCTCTCTGAAGCCTCTTACATCCCT 2277  
 QY 355 -----SerGlyProGluAspValAlaLeu 362  
 Db 2278 GCCATTCTCAAGACCACACTTTTCTCTCGCAGCGCTCTTCCGGCCCCGAGGACGTGGCTCTC 2337  
 QY 363 TyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeuLeu 382  
 Db 2338 TACATCGGCTCTGTCGCCGTGTGCTCTCATCTTGTGCTGTGCTGTGCTGTGCTGTGCT 2397  
 QY 383 ValTyrCysArgLysGlyLeuAspSerAspValAlaAspSerSerIleLeuThr 402  
 Db 2398 ATCTACTCGCGAAGAAAGGAGACTGGACTCAGATGTGGCTGACTCATCCATCTTACC 2457  
 QY 403 SerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThr 422  
 Db 2458 TCAGGCTTCCAGCTGTACAGCATCAAGCCAGCAAGAGAGACATCCCATCTGTCTCACC 2517







Web site: <http://www.nisc.nih.gov/>

Contact: nisc mgcenhgri.nih.gov  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granites, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
 Young, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E.B. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 26 Row: g Column: 22.

## FEATURES

source

Location/Qualifiers

1..2688  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4126760"  
 /tissue\_type="Brain, neuroblastoma"  
 /clone\_lib="NIH MGC\_19"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 <1..1627  
 /codon\_start=2  
 /product="Unknown (protein for IMAGE:4126760)"  
 /protein\_id="AAH09333.2"  
 /db\_xref="GI:40226528"  
 /translation="DVALYGLIAVAVCLVLLLVLYCRKXEGLDSDVADSSILT  
 SGFQVSIKSKADNPHLLTIPDLSTTTTQGSCLPRQDQSPKFLQNLHGLISPL  
 GGGTHLHSSSEAEFEVSLSTQNYFRSLPRGTSNMTYGFNLGRLMIPTNGI  
 SLLIPDAIPRGIYEIYLHKPEDVRLPLAGCCTLLSPVSCGPPGLLTRVILA  
 MDHCGEPPDSMSRLKQSCGSEWDLHGEAPSHLYXCQLEAKCYFTSQLR  
 FALVGEALSVAARLKLILPAPVACTSLEINIRVCLHDTLHDLKQVLEKQGGQ  
 LIQPRVLFKDSVHNLFSLIHDPSSLSKLIVSYQEIYFYHNGTQRYLHCTFT  
 LRVSPSTDLAKLWQVEGQGSFINFTKTRFAELALEAGVPAUVGPS  
 AFKIPFLIRQKIIISLDPKCRGADWRTLAKLHLDHLSLFFSKRSPPTAMINLWEA  
 RHPFNGNLSQLAAVAGLQDPDAGLFTVSEAC"  
 413..724  
 /note="ZUS; Region: Domain present in ZO-1 and Unc5-like  
 netrin receptors"  
 /db\_xref="CDD:smart00218"  
 1343..1594  
 /note="DEATH; Region: DEATH domain, found in proteins  
 involved in cell death (apoptosis)"  
 /db\_xref="CDD:smart00005"

## CDS

misc\_feature

misc\_feature

## ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:
2819.00	2688	
Percent Similarity:	99.63%	Matches: 540
Best Local Similarity:	99.63%	Conservative: 0
Query Match:	58.89%	Mismatches: 0
DB:	9	Indels: 2
		Gaps: 2

US-09-970-944-2 (1-899) x BC009333 (1-2688)

QY	359	AspValAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeu 378
Db	2	GAGTGGCCCTCATGTGGGCGCTCATCGCGCTGGCGCTGCTGCTGCTGCTGCTT 61
QY	379	ValLeuIleLeuValTyrCysArgGlyLeuGlyLeuAspSerAspValAlaAspSer 398
Db	62	GTCTCATCTCTGTTATTCGCCGAAGAGAGAGGCGGCTGGACTCAGATGGGTGACTCG 121
QY	399	SerIleLeuThrSerGlyPheGlnProValSerIleLeuLysProSerIleAlaAspAsnPro 418
Db	122	TCCATTCTACCTCAGCTTCAGCGCGTTCAGCATCAAGCCGACCAAGCAGAACACCCC 181
QY	419	HisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThr 437

Db	182	CATCTGCTCACCATCATCGCGGAGCCTCAGCACACACACACACACACACACACACAGTCTC 241
QY	438	CysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSer 457
Db	242	TGTCCCGGCGAGTGGGCGGAGCCCAAGTTCAGACTCACCAGTGGGACCTGTCTCAGC 301
QY	458	ProLeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGlu 477
Db	302	CCCTCGGGTGGCGGCGGCGGACACACTGCACACAGCTCTCCACCTCTCAGGCGGAGGAG 361
QY	478	PheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsn 497
Db	362	TTCGTCTCCGCGCTCTCCACCGAGAACTACTTCCCTCCCTCCCGGAGGACCGACCAAC 421
QY	498	MetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIle 517
Db	422	ATGACCTATGGACCTTCAACTTCTCGGGGCGGCTGATGATCTATACAGGAATC 481
QY	518	SerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThr 537
Db	482	AGCCTCTCTATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTACCTCAGC 541
QY	538	LeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerPro 557
Db	542	CTGCACAGCCGGAAGAGCTGAGGTTCCTCTAGCTGGCTGTTCAGACCTCTGTAGTCCC 601
QY	558	IleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAsp 577
Db	602	ATCGTGTAGCTGTGACCCCTGGCGTCTCTGTCTACCCCGCCAGTCACTCTGCTATGAC 661
QY	578	HisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysGlnSerCysGlu 597
Db	662	CACGTGGGGAGCCGAGCCCTGACAGCTGGAGCCTGGCGCTCAAAAAGCAGCTCGTGCAG 721
QY	598	GlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyr 617
Db	722	GGCAGCTGGAG---GATGCTGCCTGGCGGAGGAGGCGCCCTCCACCTCTACTTAC 778
QY	618	CysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeu 637
Db	779	TGCCAGCTGGAGGCGAGTGTCTGTCTCTTCCACCGAGCAGCTGGGCGGCTTTGCCCTG 838
QY	638	ValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaPro 657
Db	839	GTGGAGAGGCGCTCAGCGTGGCTGGCGGCAAGGCGCTCAAGCTGCTTCTGTTTGGCGCG 898
QY	658	ValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAsp 677
Db	899	GTGGCTGCACTCCCTCGAGTACCACTCCGGGTCTACTGCTGCTGATGACACCCAGAT 958
QY	678	AlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluPro 697
Db	959	GCACCTAAGAGAGTGGTGGAGAGCAGCTGGGAGAGCAGCTGGGAGAGTATCCAGAGCCA 1018
QY	698	ArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValPro 717
Db	1019	CGGCTCTCTGACCTTCAAGGAGCAGTTACCAACACCTGCGCTATCCATCCACGATGTGCC 1078
QY	718	SerSerLeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHisIle 737
Db	1079	AGTCTCTCTGTGGAGAGTAGCTTCTTGTGAGTACAGAGATCCCTTTTATCATCATC 1138
QY	738	TrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSer 757
Db	1139	TGGAATGGCAGCAGCGGTACTTTCAGCTCAGCTTCCCTTGGAGCGGTGTACCCCGCAGC 1198
QY	758	ThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPhe 777
Db	1199	ACTAGTGCAGCTGGCGCTGCAAGCTGTGGGTGTGGAGGTGGAGGCGGAGGAGAGCTTC 1258
QY	778	SerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSer 797
Db	1259	AGCATCACTTCAACATCATCCAGGAGACCAAGGTTTGTGAGCTGTGCTCTGGAGAGT 1318

QY 798 GluAlaGlyValProAlaLeuValGlyProSerAlaPheLysLeuProPheLeuLeuArg 817  
 Db 1319 GAAGCGGGGTCCAGACCCCTGGTGGGCCAGCCAGTCCCTCAAGATCCCTCTCAATTCGG 1378

QY 818 GlnLysLeuLeuSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeu 837  
 Db 1379 CAGAAGATAATTTCCAGCCCTGGAGCCACCTGTAGGCGGGTGCAGACTGCGGACTCTG 1438

QY 838 AlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThr 857  
 Db 1439 GCCCAGAAATCCACCTGGACAGCATCTCAGCTTCITGGCTCCAGCCAGCCACACA 1498

QY 858 AlaMetLeuLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeu 877  
 Db 1499 GCCATGATCTCAACCTGTGGAGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTG 1558

QY 878 AlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAla 897  
 Db 1559 GCTCGACAGTGGTGGACTGGCCAGCCAGCAGCTGGCCCTCTTCACAGTGTGGAGGCT 1618

QY 898 GluCys 899  
 Db 1619 GAGTGC 1624

RESULT 12  
 LOCUS MMU72634 9299 bp mRNA linear ROD 13-MAY-1997  
 DEFINITION Mus musculus rostral cerebellar malformation protein (rcm) mRNA,

complete cds.  
 ACCESSION U72634  
 VERSION U72634.1 GI:2088526  
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 9299)  
 Ackerman,S.L., Kozak,L.P., Przyborski,S.A., Rund,L.A., Boyer,B.B.  
 and Knowles,B.B.

TITLE The mouse rostral cerebellar malformation gene encodes an  
 UNC-5-like protein  
 JOURNAL Nature 386 (6627), 838-842 (1997)  
 MEDLINE 97271898  
 PUBMED 9126743

REFERENCE Ackerman,S.L., Kozak,L.P., Rund,L.A. and Knowles,B.B.  
 Direct Submission  
 TITLE Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street,  
 Bar Harbor, ME 04609, USA

FEATURES  
 Location/Qualifiers  
 1..9299  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57B6/SJL"  
 /db\_xref="taxon:10090"  
 /chromosome="3"  
 /map="between Amy1 and Mpmv9"

1..9299  
 /gene="rcm"  
 151..2946  
 /gene="rcm"  
 /function="transmembrane receptor"  
 /note="similar to C. elegans UNC-5"

gene  
 CDS  
 /codon\_start=1  
 /product="rostral cerebellar malformation protein"  
 /protein\_id="AA054103.1"  
 /db\_xref="GI:2088527"

/translation="WRKGLRATAACGGLGVLQMLVLPALALISASGTGSAADDE  
 FHEUPETPSPPFLPHFLPEPEAYIVKKNPNLYKASPAQCIYFKCNSEWVHO  
 KQHVDERVDEPGLFIREVSIISQVQVEELFGPEYWCQCWASSAGTITKSKAYV  
 RIAYLKTPEQPLGKEVLEQVLLQCRPEGIPVAEYVWLNKEDIIDPAEDRNFI  
 TIDHLLIIQARLSDTANYTCVAKNIVAKRKSTTATVIIVYVNGWSTWTEVSCVNSRC

## ORIGIN

Alignment Scores: 6.4e-162 Length: 9299  
 Pred. No.: 2705.50 Matches: 515  
 Percent Similarity: 73.17% Conservative: 153  
 Best Local Similarity: 56.41% Mismatches: 214  
 Query Match: 56.52% Indels: 31  
 Gaps: 12

US-09-970-944-2 (1-899) x MMU72634 (1-9299)

QY 9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28  
 Db 226 CTTGCCCTG-----GCCCTGCTAAGCGCCAGTGGCACCGGCTCCGGCGCTCAAGATGAT 279  
 QY 29 Ala-----ThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45  
 Db 280 GAATTTTTCAGAACTCCAGAAACCTTTCCATCTGACCCACCTGAGCCATTGCCACAC 339  
 QY 46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65  
 Db 340 TTCTCATTTAGCCCGGAGAAAGCTTACATTTGTGAAGAACAGCCCTGTGAACCTGTATTGT 399  
 QY 66 LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnVal 85  
 Db 400 AAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGAACAGAGAGTGGTTCATCAGAAG 459  
 QY 86 AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg 105  
 Db 460 GACCACGTAGTAGACGAGAGATAGATAAACCTCTGGTCTAATTGTGAGAGAGTGAGC 519  
 QY 106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125  
 Db 520 ATTGAGATTTTACGCCACGAGTGTGAGAACTGTTGGCCCTGAGATTAAGTGTGGTCCAG 579  
 QY 126 CysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145  
 Db 580 TGTGTGGCTGGAGCTCAGCAGGCACTACGAAGAGTTCGGAAGGCATAGCTGGCATTCG 639  
 QY 146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly 165  
 Db 640 TATCTGCGGAAGACATTCGACGAGAACCTTTGGGAAAGAGAGTGTCTCTGGAGCAGGA 699  
 QY 166 IleValLeuProCysArgProGluGlyIleProProAlaGluValIleGluTrpLeuArg 185  
 Db 700 GTCTTACTCCAGTGTCCGCCACCTGAAGGATCCAGTGGCTGAGTGGATGGCTAAG 759  
 QY 186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205  
 Db 760 AATGAAGACATAAATTGATCTCTGCTGAAGATCGAACTTTTATATATCTATCATCATCAAC 819  
 QY 206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225  
 Db 820 CTGATCATCAAGCAAGCCCGACTCTCAGATACAGCAATTAATACCTGTGTGCCAAAAAT 879  
 QY 226 IleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrp 245  
 Db 880 ATTGTGTCCAGAGAAAGAACACACGACCTGTCTATCTGTATGTATTATGTGCTGCG 939  
 QY 246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265

GRGYQKTRTCTNPAPLNGAFCEGSGVQKXIACITTLTCLPVDGRWTSKWSKNCSTCGTECH  
 WRRECTAPAKNGKDCDGLVKNKCTDGLCWAAPSDSDVALYVGIIVIAVTCLIA  
 ITTVVAFVKNRHRDFESDIIDSSALNGCFQPNVIAKARQDILLAVPDLTSSAAAYR  
 GPVYALHDVSKIPMTNSPILDEPNLIKIVYSSNAVTPQDDLAEPSSKLSPOMTOS  
 LLENEALNKNQSLQRTDPSCTAFGTFSGLGHLIIPNSGVLSLLIPAGALPQGRVYE  
 MYTVHRKNMRPFMEDSTLLTPVVSQGPGLLTRPVLTLHHCADPSTEDWKIQOL  
 KNAVQGWEDVVVVGEBENFTTCYIQDAEACHILTENLSTVALVQSTTKAAAKEL  
 KLAIFGLCCSSLEYSIRVYCLDDTDALKEVLQEROMGQQLLEPKLHFKGSIH  
 LRLSIHDIAHSLWKSLLAKYQRIPEFHIWGSQORNHLCTFTLERLSLNTVLCVKC  
 VROVEGEGQIFQNCITVSEPTGIDILPLDPASTITVTGFSAPFSLPIPIROKLCSLI  
 DAPCTRGHDMRLAHLKLNLDRLYNYFATKSPSTGVLIDLWEAQNFPGNLSMLAAVLE  
 EMGRHETVVSAAEQY"

Db 940 TCCACCTGGACAGAGTGTGTGTGTAAACGCGCTGTGGGCGAGGATATCAGAAACGC 999  
QY 266 SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsn 285  
Db 1000 ACAAGAACTGTACCAACCCAGCCCACTCAATGTGGGGCTTCTGTGAGGGGAGAGT 1059  
QY 286 ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrp 305  
Db 1060 GTG---CAGAAAATAGCATGACCTACGTATGTCCAGTGGATGGTAGTGACCTTCATGG 1116  
QY 306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325  
Db 1117 AGCAAAATGGTCAACTGTGGGACTGAATGACCCCACTGGCGCAGGAGGAGTGTACAGCA 1176  
QY 326 ProAlaProArgAsnGlyGlyGluGlyCysGlnGlyThrAspLeuAspThrArgAsnCys 345  
Db 1177 CCAGCCCCAAGAACGGGGTAAGAGCTGTGATGGCTGTGCTCCCAATCCCAAGAACTGC 1236  
QY 346 ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365  
Db 1237 ACTGATGGCTGTGCATGCAGGCTGCTCCTCACTCAGATGATGGCTCTCTACGTGGG 1296  
QY 366 Leu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeuValTyr 384  
Db 1297 ATTGTGATCGTGTAAACAGTCTGTGGCGATCACTGTGTGGTGGCCCTGTGTGTGTAT 1356  
QY 385 CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGly 404  
Db 1357 CGAAGAACCAACCGTGACTTTGAGTGTGACATCACTGACTCCTCAGCACTCAATGGCGGC 1416  
QY 405 PheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThrIleGln 424  
Db 1417 TTTCAAGCTGTGAACATCAGGCTGCCAGACAGAT-----CTCCTGGCTGTCCCC 1467  
QY 425 ProAspLeu---SerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 443  
Db 1468 CCTGACCTACCTCAGCTGAGCGCATGTACAGGGGACCTGTCTATGCTGTGCATGAT--- 1524  
QY 444 ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyArg 463  
Db 1525 GTCTCAGACAAAATCCCAATGACCACTCTCCAAATCTGACCCCACTACCAACTTGAA 1584  
QY 464 HisThrLeuHisHisSerSer-----ProThrSerGluAlaGluGluPheVal 479  
Db 1585 ATCAAAAGTGTACACACACTCAGGTGTGTCTACTCTCAGATGACCTGCGAGTTCTCA 1644  
QY 480 SerArgLeuSer-----ThrGlnAsnTyrPhe----- 488  
Db 1645 TCCAAACTGTACCCCAAGATGACCCAGCTCTGTGTAGAGAAATGAGGCCCTTAACCTGAAG 1704  
QY 489 ---ArgSerLeuProArgGlyThr-----SerAsnMetThrThrThrThrThrThrThr 505  
Db 1705 AACCAAGCCTCGCAAGACAGACTGACCCATCTCTGACAGCATTTGGTACCTTCAACTCT 1764  
QY 506 LeuGlyGlyArgLeuMetIleProAsnThrGlyLysSerLeuLeuLeuLeuProAspAla 525  
Db 1765 CTGGGGGTCACTCATCTCTCTATTCAGAGTAGCTGTCTGATTCCTCCGCTGGGACC 1824  
QY 526 IleProArgGlyLysIleThrGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 545  
Db 1825 ATTCTTCAGGGAGAGTCTATGAATGTATGTGACTGTACACAGGAAAGAAATATGAGG 1884  
QY 546 LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly 565  
Db 1885 CCCCCATGAAGACTCTCAGACCCCTACTTACCCCTGTGTGTGAGCTGTGGCCCTCTCTGA 1944  
QY 566 ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp 585  
Db 1945 GCTCTGTGACCGCCCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 604  
QY 586 SerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeu 605

Db 2005 GACTGGAGATCCAGCTCAAAAACAGGACAGGACAGGACAAATGGGAG---GATGTTGTG 2061  
QY 606 HisLeuGlyGluGluAlaProSerHisLeuTyrTrpCysGlnLeuGluAlaSerAlaCys 625  
Db 2062 GTGGTGTGGGAGGAGAACTTCCAAACCCCTGTTACATTCAGCTGGATGCGAGGCTTGC 2121  
QY 626 TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAla 645  
Db 2122 CATATCTTCACAGAAACCTCAGTACCTATGCTTGGTGGGCGAGTCCACCAACAGCA 2181  
QY 646 AlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr 665  
Db 2182 GCTGCCAAGCGTCTTAACTGGCCATCTTGGGCGCCCTCTCTGCTGCTCTCTCTGGAGTAC 2241  
QY 666 AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu 685  
Db 2242 AGCATTAGAGTCTACTCGCTGGATGCACACAGCATGCCCTGAAGGAAGTTCTCAACTG 2301  
QY 686 GluLysGlnLeuGlyGlyGlnLeuLeuLeuGlnProArgValLeuHisPheLysAspSer 705  
Db 2302 GAGAGGCAATGGGAGGACAGCTCTTAGAAGAACCCCAAGGCTCTTCATTTTAAAGGCAGC 2361  
QY 706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeu 725  
Db 2362 ATCCACACCTGCGCTGTCTATTCATGACATGCCCATCTGGAGTGGCTCTCAAGAACTC 2421  
QY 726 LeuValSerTyrGlnGluLeuProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeu 745  
Db 2422 CTGGCTAAGTATCAGGAATTTCCATTTTACCATCTGGAGTGGCTCTCAAGAACTC 2481  
QY 746 HisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeu 765  
Db 2482 CACTGCACCTTCACTCTGGAAGACACTCAGCTTAAACACAGCTGGAAGTGTTCGAACTC 2541  
QY 766 TrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLys 785  
Db 2542 TGTGTGGCGGAGGTTGAGGAGGAGGACATCTTCCAGCTCACTGCTGTGTGAGAG 2601  
QY 786 AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuVal 805  
Db 2602 GAACCTACTGGCATCGACTTACCTCTCTGGACCTCTAGTACCATCACCAGCTGCACC 2661  
QY 806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825  
Db 2662 GGACCAAGTCTTTCAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2721  
QY 826 ProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845  
Db 2722 GCCCTCAACACAGAGGCGCATGACTGGAGATGCTGGCCCATAACTCAACCTGGACAG 2781  
QY 846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGlu 865  
Db 2782 TACTTGAATTAATTTGCCACCAATCGAGCCCACTGGCGTAATCTCTGGATCTTTGGAA 2841  
QY 866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGly 885  
Db 2842 GCACAGAACTTCCAGAGAAACCTGAGCATGCTGGCAGCGCTCTCTGGAAGAAATGGGA 2901  
QY 886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898  
Db 2902 AGACATGACACAGTGTGTGCTTGGCAGCAGGAAGGACAG 2940

RESULT 13  
BD057525  
LOCUS  
DEFINITION  
Netrin receptors.  
BD057525  
ACCESSION  
BD057525.1 GI:22603131  
VERSION  
JP 2001505062-A/2.  
KEYWORDS  
synthetic construct  
SOURCE  
ORGANISM  
artificial sequences.  
REFERENCE  
1 (bases 1 to 1787)



LOCUS AB118026 9328 bp mRNA linear ROD 19-DEC-2003  
 DEFINITION Rattus norvegicus Unc5h3 mRNA, complete cds.  
 ACCESSION AB118026  
 VERSION AB118026.1 GI:40217509  
 KEYWORDS  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1  
 Kuramoto, T., Kuwamura, M. and Serikawa, T.  
 Rat neurological mutations cerebellar vermis defect and hobble are  
 caused by mutations in the netrin-1 receptor gene Unc5h3  
 Mol. Brain Res. (2003) In press  
 2 (bases 1 to 9328)  
 Kuramoto, T. and Serikawa, T.  
 Direct Submission  
 Submitted (22-AUG-2003) Takashi Kuramoto, Institute of Laboratory  
 Animals, Graduate School of Medicine, Kyoto University;  
 Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan  
 (E-mail: tkuramoto@anim.med.kyoto-u.ac.jp)  
 URL: www.anim.med.kyoto-u.ac.jp, Tel: 81-75-753-4494,  
 Fax: 81-75-753-4409  
 FEATURES  
 source  
 1..9328  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /chromosome="2"  
 gene  
 1..9328  
 /gene="Unc5h3"  
 151..2946  
 /gene="Unc5h3"  
 /codon\_start=1  
 /protein\_id="BA005191.1"  
 /db\_xref="GI:40217510"  
 /translation="MRKGLRATAPAAAGLGLGLYLQMLVLPALALLSASGTGSAODDD  
 FFHELPETPSPDPPELPPLHPIEPERAYIVKNKPNVLYCKASPATOIYFKCNSEWVHQ  
 KHWDERVDTEGLTVREVSIEISQOVELFGPEDYWCQCVANASAGTTKSKAYV  
 RIALLKTEQEPGLGEVSLQEVLLQCRPEGLPEMAEVEWLNEDIIDVEDRNFYI  
 TIDHNLILKQARLSDTYICVAKNIVAKRKSTTATVIVYNGGWSWAEVSCNCR  
 GRGQRKTRCTNPALNGAFCEGQSVQKIACTILCPVDGRWTSKWSKSTCTGTECH  
 WRRECTAPKNGKDCDGLVLQSKNCTDGLCQAPSDDDVALVGVIVAVTVCLA  
 ITVVVALFVYRKNRHPFESNIIDSALNGGFQPVNKAARQDILAYPPDLTSAAMYR  
 GPVVALHDVSDKIPMNTSPILDPNLKIKVYNSGNVTPQDDLAFESKLSFQMTQS  
 LLENEALNKKQSLARQTPSCFTAGTFNLSGLGHLLIPNSGSLIPAGAIPOGRVYE  
 MYTVIRKENRPFMDSSQLLPVSVCGPPGALLTRPVILTLHHCADPENTEDWKQL  
 KNAVQGWEDVVVVEENFTPCYIQLDAEACHILTENLSTVALVQSQTKAAAKRL  
 KLAIFGLCCSLEYIRVYCLDQDALKVELQLEROMGQGLLEPKALHFKGSITH  
 LRSLIHDITSLWKSLLAKYQEI PFYHWSGSORNLHCTFTLERISLNTVELVCKL  
 VRQVEGEGQIFOLNCTVSEPTGIDLPILDPASTITVTGPSAFSPILPIRQKLSCL  
 DAPOTRGHDWRMLAHKLNLDRLYNLPATKSSPTGVLLDLEAQNFPDGNLSMLAAVLE  
 ENGRHETVSLAAGGY"  
 2947..9328  
 /gene="Unc5h3"  
 8727..8772  
 /note="simple sequence repeats"  
 /rpt\_unit="gt"  
 9309..9314  
 /gene="Unc5h3"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,74e-161 Length: 9328  
 Score: 2688.50 Matches: 512  
 Percent Similarity: 72.95% Conservative: 154  
 Best Local Similarity: 56.08% Mismatches: 216  
 Query Match: 56.16% Indels: 31  
 DB: 10 Gaps: 12  
 US-09-970-944-2 (1-899) x AB118026 (1-9328)

QY 9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28  
 |||||  
 Db 226 CCTGCCCTG-----GCCCTGCTCAGCGCCAGTGCACCGCGCTCGGCCGCTCAAGATGAT 279  
 |||||  
 QY 29 -----AlaThrValAlaAsnProValProGlyAlaAsnProGlyAlaAsnProAspLeuLeuProHis 45  
 |||||  
 Db 280 GACTTTTTCATGAACCTCCAGAAACCTTCCATCAGACCCACCTGAGCCATTCACACAC 339  
 |||||  
 QY 46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65  
 |||||  
 Db 340 TTTCTCATTTGAGCCCGAGGAAGCTTACATTGTGAAGAACAAGCCTGTGAACTGTATTGT 399  
 |||||  
 QY 66 LysAlaValProAlaThrGlnIlePhePhePheLysCysAsnGlyGluTrpValArgGlnVal 85  
 |||||  
 Db 400 AAGCCAGCCTGCGCCACCATCTTCTCAAGTGCACACGCAATGGGTTTCATCAGAAG 459  
 |||||  
 QY 86 AspHisValIleGluArgSerThrAspGlySerThrGlyGluProThrMetGluValArg 105  
 |||||  
 Db 460 GACCACGTAGTAGAGAGAGAGGTAGTAGAAACCTCCCGTCTCATGTTGAGAGAAAGTAGAGC 519  
 |||||  
 QY 106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125  
 |||||  
 Db 520 ATTGAGATTTCAGCCAGCAGGTGGAGAACTGTTCGGCCTGAAGATTACTGTGTCCAG 579  
 |||||  
 QY 126 CysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145  
 |||||  
 Db 580 TGTGTGCCCTGGAGCTCAGCGGCGACTACGAAAGTCCGAAGGCATACGTGCGCATTCGC 639  
 |||||  
 QY 146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGly 165  
 |||||  
 Db 640 TATCTCGGAGACATTTGAGCAGGACCTTGGGAAAGAGTGTCTTGGAGCAGAA 699  
 |||||  
 QY 166 IleValLeuProCysArgProGluGlyIleProProAlaGluValGluTrpLeuArg 185  
 |||||  
 Db 700 GTGTTACTCCAGTCCGCGCCACCTGAAGGATCCCAATGCTCGAGGTGGATGGCTAAAG 759  
 |||||  
 QY 186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205  
 |||||  
 Db 760 AATGAGACATAATTGATCTCTGCGAAGATCGGAATTTTATATCATATCATGATCACAAC 819  
 |||||  
 QY 206 LeuValAlaArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225  
 |||||  
 Db 820 CTGATCATCAAGCAGCAGCAGACTCTCAGATACAGCAATATATACCTGTGTGCCAAAAAC 879  
 |||||  
 QY 226 IleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAenGlyGlyTyr 245  
 |||||  
 Db 880 ATTTGTTGCCAAGAGGAAAAAGCACCACAGCCACCGTCATCGTGTATGTTAATGTTGCTGG 939  
 |||||  
 QY 246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265  
 |||||  
 Db 940 TCACCTGGGCGAATGGTCTGTGTGTAAACAGCCGCTGTGGCGAGGATATCAGAAACGC 999  
 |||||  
 QY 266 SerArgSerCysThrAsnProAlaProLeuAsnGlyValAlaPheCysGluGlyGlnAsn 285  
 |||||  
 Db 1000 ACAAGAACCTGCACCAACCCGCCCTCATGTTGGGGCTTCTGCGAGGGCGAGAT 1059  
 |||||  
 QY 286 ValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrp 305  
 |||||  
 Db 1060 GTA---CAGAAAATAGCATGCACCGTTATGCCAGTGCATGGCAGGTGGACTTCGTGG 1116  
 |||||  
 QY 306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325  
 |||||  
 Db 1117 AGCAAAATGCTCAACTGTGGGACTGAGTGACCCACTGGCGGAGGAGGAGGTACAGCA 1176  
 |||||  
 QY 326 ProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys 345  
 |||||  
 Db 1177 CCAGCCCCCAAGAACGGGGTAGAGCTGTGATGCGCTGCTCCTCAATCCAAAGAACTGC 1236  
 |||||  
 QY 346 ThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyrValGly 365  
 |||||  
 Db 1237 ACCGATGGGCTGTGCATGCAGGCTGCTCTCTGACTCAGATGACGTGCTCTCTACGTTGG 1296  
 |||||  
 QY 366 Leu---IleAlaValAlaValCysLeuValLeuLeuLeuLeuValTyr 384



Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Wyman,B., Wu,X., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Triggiano,J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

# JOURNAL

## REFERENCE

Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 200968)

Biren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ranasingh,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,B., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

# JOURNAL

## COMMENT

Submitted (13-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 13, 2003 this sequence version replaced gi:28201591.  
All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

## Project Information

Center project name: L26431

Center clone name: 335\_A.11

## Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 136489 bases at least Q40  
Consensus quality: 199303 bases at least Q30  
Consensus quality: 199583 bases at least Q20  
Insert size: 197000; agarose-fp  
Insert size: 200168; sum-of-coverage  
Quality coverage: 9.9 in Q20 bases; agarose-fp  
Quality coverage: 9.7 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 44442: contig of 44442 bp in length  
\* 44443: gap of 100 bp  
\* 44543: contig of 697 bp in length  
\* 45239: gap of 100 bp  
\* 45339: gap of 1128 bp in length  
\* 45340: contig of 100 bp  
\* 46468: gap of 100 bp  
\* 46567: contig of 8951 bp in length  
\* 55518: gap of 100 bp  
\* 55619: gap of 100 bp  
\* 122673: contig of 67055 bp in length  
\* 122773: gap of 100 bp

\* 122774 137362: contig of 14589 bp in length  
\* 137363 137462: gap of 100 bp  
\* 137463 162118: contig of 24656 bp in length  
\* 162119 162218: gap of 100 bp  
\* 162219 187047: contig of 24829 bp in length  
\* 187048 187147: gap of 100 bp  
\* 187148 200968: contig of 13821 bp in length.

## FEATURES

source  
1. .200968  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-335A11"  
/clone\_lib="RPC1-23 Female Mouse BAC"  
1. .44442  
/note="assembly\_fragment"  
clone end:SP6  
vector\_side:left  
44543..45239  
/note="assembly\_fragment"  
45340..46467  
/note="assembly\_fragment"  
46568..55518  
/note="assembly\_fragment"  
55619..122673  
/note="assembly\_fragment"  
122774..137362  
/note="assembly\_fragment"  
137463..162118  
/note="assembly\_fragment"  
162219..187047  
/note="assembly\_fragment"  
187148..200968  
/note="assembly\_fragment"  
clone end:T7  
vector\_side:right

## ORIGIN

Alignment Scores:  
Pred. No.: 1.27e-158 Length: 200968  
Score: 2681.00 Matches: 771  
Percent Similarity: 25.66% Conservativeness: 28  
Best Local Similarity: 24.76% Mismatches: 42  
Query Match: 56.01% Indels: 2278  
DB: 2 Gaps: 18

US-09-970-944-2 (1-899) x AC123700 (1-200968)

QY 59 LysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePhe----- 75  
Db 101981 AGCCATACCTGCTTACATCCAGAGATAAACACGACGATCCCTGGAGTTCCCTCCAGGC 101922  
QY 76 -----LysCysAsnGly-----GluTrpValArgGlnValAspHisVal 88  
Db 101921 CTCGGAGGCTACAGGGGCTATAAGCTCTAGTGGTGGGGGGGCAATCAACCAATTG 101862  
QY 89 IleGluArgSerThrAsp-----GlySerSerGlyGluProThrMetGluValArg 105  
Db 101861 GGACGGGAGACTCAGTGTGGCTCTGAACCCAGCAGGATTGCCAACCATGGAGGTCCGG 101802  
QY 106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTrpCysGln 125  
Db 101801 ATCAAGTATCAAGGCAGCAGGTCGAGAAAGTGTGGCTGGAGGAGTACTGGTCCAG 101742  
QY 126 CysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTrpIleArgIleAla 145  
Db 101741 TGTGTGGCATGGAGGCTCTCTCAGAACCAACCAAGAGCCAGAGGCTTACATCCGATTGCC 101682  
QY 145 ----- 145  
Db 101681 TGTGAGTACTGGACCAGGCTGTGGCATCCGGGCGATGCTACTAGAGGGTATCCCTCGGA 101622  
QY 145 ----- 145



Db	101621	GAGAGGGCGGCAAGCTTCCTGGAGGGTAGAGGCAGTGCCTTAGGAACATAAATCACCCAC	101562	Db	100543	GAGCAGCCCACTCACCCGAGCTTATAAGAACTTTGCCCTCTCTGCCCATCCCTTGGCCC	100484
QY	145	-----	145	QY	242	-----	Asn 242
Db	101561	ACTCCTCAATGCACATGGTGCCTCCTTCGGCCATGACGTGTGCATGACGTCTGCAGGC	101502	Db	100483	AGCATCTCTAGAGAACCCCGCTCCTCGACCCAGTCTCTCTCTGTCCGCCAGTGAAC	100424
QY	146	-----	ArgLeuArgLysAsnPheGlu 152	QY	243	GlyGlyTyrSerThrThrGluTyrSerValCysSerAlaSerCysGlyArgGlyTyr	262
Db	101501	CAAGAGTCAGTCATCCAGACCCCGTTTCTCTCCAGATTTCGCAAGAACTTTGAG	101442	Db	100423	GGTGGGTGTCGACGTGGACCGAGTGGTGGTCTGCAGTCCAGCTGTGGCGGTGGCTGG	100364
QY	153	GlnGluProLeuAlaLysGluValSerLeuGlnGlnGlyLeuValLeuProCysArgPro	172	QY	263	GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyAlaPheCysGlu	282
Db	101441	CAGAGCGCTGGCCAAAGGAGTGTCACTGGAGCAAGCATTTGCTTACCTTTGCGCCC	101382	Db	100363	CAGAAACGAGCGAGCTGCACCAACCCGCGACCTCTCAACGGGGCGCTTCTGTGAG	100304
QY	173	ProGluGlyLeuProProAlaGluVal	181	QY	283	GlyGluAsnVal	286
Db	101381	CCGGAGGAATCCCCCAGCTGAGGT-GAGTGGAGGGTGGGGCTTCCAGACCTCTATC	101323	Db	100303	GGCGAGAATCTCCAGAAAACAGCTGGGCCACTCTGTGCCAGGTACCGGCAGCCACCGA	100244
QY	181	-----	181	QY	286	-----	286
Db	101322	CTACTAAAGTCAGACTCCTGGTGCACCTCCCGAGTGGTGCCTCTATACCTGTTC	101263	Db	100243	CTCTGTCATCGTTCTCAACACGCCATCTCCGTGGCTGGCTCCATCGCACCCACACAGCC	100184
QY	182	-----	GluTyrLeuArg 185	QY	286	-----	286
Db	101262	GGCCAGCAGCGGCGCAGCGGTGTGTGTACCGCTTCTCCCGCTAGGTGGAGTGGCTCGA	101203	Db	100183	TGCCCGCCCTCCCCCATGGCTCCATCCACCCGCTGCCACACACAGGCCACAGGTCA	100124
QY	186	AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer	205	QY	286	-----	286
Db	101202	AATGAGACCTCTGTGGACCTCTCCCTCGACCCCAATGTGTACATCACACGGGAGCAGAC	101143	Db	100123	GTCGTAGGTGGGATCAGGGTGCAGTGTGTGCAGTCAAGGGCTCGGTTAGGGTCTAA	100064
QY	206	LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn	225	QY	286	-----	286
Db	101142	CTAGTCTGCGCAGCGCGCTGCGGCACACTCCCAACTACCTCGCTGGCGCAAGAAC	101083	Db	100063	GGTCAGTTTTAGSACAGGATCTAGGCTCGCGCTGGGGCTGCACAATCCACCTGTGCCTG	100004
QY	226	IleValAlaArgArgArgSerAlaSerAlaAlaValIleVal	241	QY	286	-----	286
Db	101082	ATCGTGGCCCGTGGCGAAGCGCTCTCGCGCCGTCATTGTTATGG-TGGGTCTGGGC	101024	Db	100003	GACCAGAAGGCCCATCATGTACTAGGAGCCCTTACCGCCTCAGACAGGGACAGCCTCC	99944
QY	241	-----	241	QY	286	-----	286
Db	101023	TCCCAGTCACAGGAAGAGGTACTGAGGTGCCAACAGGACTGTGTGTGGCTCAGAGG	100964	Db	99943	CTGGTTGAGAGAGCCGAGCAGGCCCCCAGCATCATTTGGTCTCTTGTTCATTAGCATGTG	99884
QY	241	-----	241	QY	287	-----	HisAspArgThrValSerSer
Db	100963	ACGGAAGGCATGGAGCAAGAACAGGCTAACAGACAGTGGCTTGACTAGAAAGAG	100904	Db	99883	TCATCAAGGCTAGCTGGACACTCAAGTCCGACCGATCAGAGGGCTCACCTTGGCATTTG	99824
QY	241	-----	241	QY	293	-----	293
Db	100903	CAGAAAGTGTATCTGGGCAAGGATGGACATGTGGTCACTTGTGAGTGGGAGCATGG	100844	Db	99823	TACCCAGCCAGGCGCTGTGTGTCCAGATGACTAGCGGGGCCACCCAGGCCAGCCACC	99764
QY	241	-----	241	QY	294	-----	LeuLeuValSerVal 298
Db	100843	TGCAGAGAGCAAGAGACAGAGAGGAGGCTCAGAGGAGAGTAACATCAACCTTGGCGGT	100784	Db	99763	CAGGCCAACCTCTTAAGCTGACTGTCCCATTTCCCTCTCTGCCTGTCTCTTGTTCAGTG	99704
QY	241	-----	241	QY	299	AspGlySerThrProTyrSerThrSerAlaCysGlyLeuAspCysThrHisTrp	318
Db	100783	CTCAGTCCCGAAATGGGAGAGAGCGCTCAGAAATGAGACGCTTCTCCCTCCCAAGAGCA	100724	Db	99703	GATGGAGCTGGAGCCCATGAGTAAGTGGTCAGCTCGGGCTTGAAGTCCACCCACTGG	99644
QY	241	-----	241	QY	319	ArgSerArgGluCysSerAspProAlaProArgAsnGlyGluGluCysGlnGlyThr	338
Db	100723	GAAGAGGTGGATAGGAGCTGCCCTGCGAGAAGCTGGGTTCTAGCTGGAGGTGCCCTGT	100664	Db	99643	CGAGCCGGGAGTGTCTCGACCCCGCCCGAACGAGGAGTGGAGTGGCGGGGTGCT	99584
QY	241	-----	241	QY	339	AspLeuAspThrArgAsnCysThrSerAspLeuCysValHis	352
Db	100663	GACCTGTAGTTTAGTCTGGCCACCTTGACACTTGGCCAGAGTAGATATCTGAAGGCTCT	100604	Db	99583	GACCTGGACACCCGCAACTGTACAGTACCTCTGCTGACACAGTGAAGTCTCTCTGCC	99524
QY	241	-----	241	QY	352	-----	352
Db	100603	GCCAGCCCTGGGGCCAGCAGCCTAGGCTTGGTCTCTCAGGGTTTCTCCAGGCACAG	100544	Db	99523	CAGGGCTCACCTCCGCTCTGACAGCCCTATGGGCTGTGGTGTAGGAGCCACTCTCTGCC	99464
QY	241	-----	241	QY	352	-----	352
Db	100544	-----	241	Db	99463	TGTGTCTCTCCAAAGCTTGGGAGCCAGGGGTCCCTTTTCATCTCTAGCACACTCTTCCCT	99404



QY 352 ----- 352  
Db 99403 GTTACACAGACACACCCTTCCCTCCTGAGACATGCGATGTGAATCTTGACCCCTGT 99344  
QY 352 ----- 352  
Db 99343 ACATCCAACTGCTCCCGCGGCTTTGAAATTATCCCGTTCCCAAGCTCACACATAAGT 99284  
QY 352 ----- 352  
Db 99283 CCTGTGGGGGTGCTGGGCTTGTGAGGCTGTGTAATTACAAGGCCATAACTAACAATTAAA 99224  
QY 352 ----- 352  
Db 99223 GATCTCTGTTGGAATGTTTTTCTTTAGACTACACCTCCCTGTTCTTTCAATTAAA 99164  
QY 352 ----- 352  
Db 99163 AACCTAATTATCTGTGTCATTCCTTTCTGATGAGACATTTCTTCCATTTATGAAAACTCT 99104  
QY 352 ----- 352  
Db 99103 TAAAAATGATAGTCTTCTCCAAGCAGTTGGTCCCCACCACCCAGTTTCAGGGCT 99044  
QY 352 ----- 352  
Db 99043 GGTGGTATACACAGGGCTCACACTGTCTGGAACACCTGAGAACTGACCACCCATGT 98984  
QY 352 ----- 352  
Db 98983 ATCTGCTATTCTTCTATTCTATAATGAAGTACTAGCAAGAAATTTATTAGTTTATAGAGC 98924  
QY 352 ----- 352  
Db 98923 TCAAATTCAAATAGCTGATGTGGCTCTGGTAAATGTGCTTTCTACCCCAAGCCCAT 98864  
QY 352 ----- 352  
Db 98863 TTCTGTGTGCAGATAGCACCATGTCTAGTACTGTAGACAGAGAGATCACGTGTGAGG 98804  
QY 352 ----- 352  
Db 98803 AAGAGAAAAAGACAGACACAGACAGTAGACACGGAAGATGGACAGCCCATGTGTCT 98744  
QY 352 ----- 352  
Db 98743 AAGAGCAGCACTCATATGATGTAAATGCTCTCCCTCCCTGAGGCCCCACCTCTAAAAGGCC 98684  
QY 352 ----- 352  
Db 98683 CACTGTCTCATATTTGCCAATGGGGACACACCTGTCAATATGATCCTTTAGGCAGG 98624  
QY 352 ----- 352  
Db 98623 CATGTCAAACTATAGCACCATCTTAGGGTCTCTCAGGAGTTTACCTTGCCCCAGACAC 98564  
QY 352 ----- 352  
Db 98563 CCTGACCCCGAGCTAGGAGCCAGGATATACCACTCACCCCGGTGAAGTACTCTT 98504  
QY 352 ----- 352  
Db 98503 CCCCTGCTCCACTGTGCTGTGAGGTGTGACTTGTCTCAATCCCAAGCAAGTTGTGGT 98444  
QY 352 ----- 352  
Db 98443 GGTATAGAACTAGGATGTACCCACCAATACCAAGCTCCCTGCGCAGGTGCCACTGCCAG 98384  
QY 352 ----- 352  
Db 98383 GGGGTGCTGAGATCCTATTCTCTGTGGGATAAACCATGGGCTCTCTCTGAAGCCTCTA 98324

QY 353 -----Ser-AlaSerGlyProGluAspVa 360  
Db 98323 CATCCCTGCCANTCTCAAGACACACTTTCTTTGACAGCTCTTCCGGCCCCGAGAGCGT 98264  
QY 360 lAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValIle 380  
Db 98263 GGCTCTCTACATCGCCCTCGTCCGCGTGGCTGCTCATCTTGTCTGTCTGTCTGTCT 98204  
QY 380 uIleLeuValTyrCysArgLysGluGlyLeuAspSerValAlaAspSerSerIle 400  
Db 98203 CGTCTCATCTACTCTCCGCAAGAGAGACTGGACTCAGAGTGGCTGACTCATCCAT 98144  
QY 400 eLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla 415  
Db 98143 CCTTACCTCAGGCTTCCAGCCTGTGAGCATCAAGCCAGCAAGCAGGTGAGGTGCCACC 98084  
QY 415 ----- 415  
Db 98083 TACCCCTGCTCTCCTGCTGAGCACCCCAAGCAAGGCTGTTGATGGGGAGTCGCTA 98024  
QY 416 -----AspAsnProHisLeuThrIleGlnProAspLeuS 428  
Db 98023 CCTGAATCCATCTTTATCTCTGTAGACAATCCCATCTGCTCACCATCCACCGGACCTCA 97964  
QY 428 er---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProL 447  
Db 97963 GCACCACGACACCATCTACAGGGCAGCCTGTCTCCCGCAGAGTGGACCCAGCCCA 97904  
QY 447 ysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuH 467  
Db 97903 AGTTCAGCTCTCTAATGGTCACTGCTGAGCCACTGGGAGTGGCGCCCATAGCTGC 97844  
QY 467 ishSerSerProThrSerGluAlaGluPheValSerArgLeuSerThrGlnAsnT 487  
Db 97843 ACCACAGTCCCCACCTCTGAGGCTGAGGACTCGTCTCCGCTCTCCACCCAAAAC 97784  
QY 487 yrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuG 507  
Db 97783 ACTTTCTGCTCTGCCCCGGGTACCGACCAACATGGCCTATGGGACCTTCAACTTCTCG 97724  
QY 507 lyGlyArgLeuMetIleProAsn----- 514  
Db 97723 GGGGCCGCTGATGATCCCTTAAACAGGTAGGCCCCCTAGAACTCCCGAGAGAACATGAGA 97664  
QY 514 ----- 514  
Db 97663 GGTAGACTAGTCTTGTCTATCTTTTCTGCCAGAGTGCACCGCTCGTGCAGAGTGGCC 97604  
QY 514 ----- 514  
Db 97603 CAGTCTGGCCAGTATAGGAAGGATGAGTTGGCGCTATAGCAAGCTGTGAAAGGGTC 97544  
QY 514 ----- 514  
Db 97543 TCCTTGTGTGTGGACACGAAAGGATTGTATGTTCAITTTATCACACAGTACGCTTAC 97484  
QY 514 ----- 514  
Db 97483 ATGTTGATAGGTGCTGATCGATCGGCCCCCTAAACACAGACCCAGAGGCTCACAAAACAAGAC 97424  
QY 514 ----- 514  
Db 97423 AGCAGAGCACCTTTAGAACCTGCCCCCTCCAGAGCTCTCCATCTCATCCACCACTCCA 97364  
QY 514 ----- 514  
Db 97363 CAGAAAGTGCCTCAGACAGGGTCCCTGAGACCTTCTTGTAGAGATCCACAGACCTCCCA 97304  
QY 514 ----- 514  
Db 97303 AGGCCATCTTGGAGAAATCTCTGCAACCTCTGACAAGGCTGACCAAGGCTGACCCCACTCTGACAT 97244  
QY 514 ----- 514

Db	97243	TCCTGTTCCTTCCTT	AGGTGACCCCA	CATATCTCCTGTTCCTGTCCTATCCCTGGCTGT	97184
QY	514	-----	-----	-----	514
Db	97183	TCCGGTATAGCGTTCCTC	TAGGATACCCCTCCCA	CAACCCACCCCACTTCGCGCCGTGC	97124
QY	514	-----	-----	-----	514
Db	97123	CAITTCAGTGATCTCACCT	TACTCTCTCCATCCAGCCGCCAGTTCCTG	CAGACATGG	97064
QY	514	-----	-----	-----	514
Db	97063	AACTCTGCAITCAGCTGCCT	CTTGGACAGCTGCATCTGCTCTC	AGACCTCGGAG	97004
QY	514	-----	-----	-----	514
Db	97003	TCITTTGGCCGAAACAGAAC	CGGTCCTCCACCCCTTCTCT	CAGCTACCTCCTCCTGGCT	96944
QY	514	-----	-----	-----	514
Db	96943	TTCCCCGGCTCACGGTTGG	CACAGTGTCAITTTAAGCTAG	AACTGGCTGGGTAC	96884
QY	514	-----	-----	-----	514
Db	96883	CTTGGGCTCCTGTCTTTCT	ATTCCTATCCACCCCTCGCT	TATTAAGTCAGCTGTATC	96824
QY	514	-----	-----	-----	514
Db	96823	GTTGCATGCTCCCTCTCC	CAATAGCTCCTGTCTCCT	CTCCTCATCCTGCCCTGGT	96764
QY	514	-----	-----	-----	514
Db	96763	CGCTCACACACCCAGGCT	AGCTCCTCTTACCTCAAC	CGCGCATCTACACATTTACT	96704
QY	514	-----	-----	-----	514
Db	96703	GAATGAATCTAGAGTTG	ATTATGCCCCCTCCCACT	CAATCCTTCTAAGTTCCAGTT	96644
QY	514	-----	-----	-----	514
Db	96643	GCTCTTACAGAGCCCGAG	TGACAAAGGTAATGCA	ACCCCACTGCTGCTAGCCGCT	96584
QY	514	-----	-----	-----	514
Db	96583	GTGTGCTTTTCCCTCCT	ATATGCAAGTTCTGTAT	GTGACAGCCCTTACAGACCTTTGGCAC	96524
QY	514	-----	-----	-----	514
Db	96523	CTGAACATGACACTGTCC	CTGGCTCATAAAGTAAT	GGTCTTTTCTACCTTTCTCCTGGC	96464
QY	514	-----	-----	-----	514
Db	96463	TCATCTGTTGTGCTGCT	GTGCCTCATACCCCTC	AGGACAGGGCATGAGAAATCTC	96404
QY	514	-----	-----	-----	514
Db	96403	ATAGCTAGTGAAGTGAG	CTGGTCCCTGTTAGCTGT	ATAAAGGACAGATGGTGG	96344
QY	514	-----	-----	-----	514
Db	96343	GTATTCCTGGCCCTGC	AGTGTGCCACCCCTGT	GTCCTCTGTGTCTAGTTGCT	96284
QY	514	-----	-----	-----	514
Db	96283	CTCTCCTTGTGACTG	AGAAAGCTAGGCCCCCTTTT	CTCTTCTAGCTTTCTGAGTGCCA	96224
QY	514	-----	-----	-----	514
Db	96223	GCATGAGTATTCAATCT	TGTAATPACTCTCCAG	CAGGCAITTCACACGAGGAGTGTGTG	96164
QY	514	-----	-----	-----	514

  

Db	96163	ACTGCCCCGACTCCAGG	CAGCATCAGTTCTAC	GGTTCAGCACAGGTTAGTGAAGGAGC	96104
QY	514	-----	-----	-----	514
Db	96103	AAAAGATTGAGCACAT	GGTTCGAGCTGTCT	GTGAGACAGCTCGGTGGTAAGACCCATC	96044
QY	514	-----	-----	-----	514
Db	96043	ACTCGGTGTTAAGAC	ACGGACAGGTACAT	CTGTAGAGCACATTAAACCACTCGACAGC	95984
QY	514	-----	-----	-----	514
Db	95983	TTCAAGTGTGGGAGG	CTGTGTGACCTAT	TGGAACCATCACTGTGAGTTTGTAGAATATAA	95924
QY	514	-----	-----	-----	514
Db	95923	GTGTCTTACATAGCC	CTTGAACAGCTGT	AGAGTCCAGCCACGAAACAGTGGAGCTATC	95864
QY	514	-----	-----	-----	514
Db	95863	TGCTGTTTTAGGGCA	AAAGTTCAGAAATGT	CTCCCCCTTTCTCCTCCCTCAGAGCAGGACC	95804
QY	514	-----	-----	-----	514
Db	95803	TCAATGCTGCTCTCC	GTGTGTCTCAGCC	CTGTGTCTCTCCACTACAAAACCTCCC	95744
QY	514	-----	-----	-----	514
Db	95743	CTTCTCCACACACT	GATACCACTCTTGGG	GCAGAGCGCATCAACAGACCCAGAGGACA	95684
QY	514	-----	-----	-----	514
Db	95683	TTGGAGCGGTAGAT	CTTTGGCCCTTGAGG	GGCTATTATAGACTGCCCCCATCCCGCAAG	95624
QY	514	-----	-----	-----	514
Db	95623	GATGGAGACAATG	ACAGGAGAGACCAGT	TCTTAGAGGGGACAGGACTTAAGGGGGGAGGA	95564
QY	514	-----	-----	-----	514
Db	95563	GCTTGAACAAGACA	AAAGGCCCTGTCT	TGGAACCGGTGAACGTAGCAAAAGTTGGGCGA	95504
QY	514	-----	-----	-----	514
Db	95503	GGGACCGACTGCTGG	GATGACAGTCAAGAG	CCCTTCATGGCAGGCTTGGAGATGGGGAA	95444
QY	514	-----	-----	-----	514
Db	95443	GGTTTTAAGCAGAA	AGTACCTGTGTC	CCAAAGAACGGAACAGGATGTAGGACGCACCTTC	95384
QY	514	-----	-----	-----	514
Db	95383	TTGGTGTCTTCTG	CCCTACTTTTAGCTT	CTTAGCTGGGTGGCTCAGATGTTCTTGGAGTC	95324
QY	514	-----	-----	-----	514
Db	95323	TCCCTGACTCAGAGT	CTGTGCCTATGTGG	TGGTGGACAGTATGGCAGTCTTGAACACC	95264
QY	514	-----	-----	-----	514
Db	95263	TCTGTCCCAACACA	AGGGCGTCTGAAG	AGAACTCTGTGGGTCTGGCTAGTTCCTGGGTA	95204
QY	515	-----	-----	-----	520
Db	95203	GTATTGGGTATCTG	TAGCCACAAGCCGT	GACCCCTCTGCTTTTACAGGAATCAGCCTCC	95144
QY	520	-----	-----	-----	540
Db	95143	TCATACCCCGGAC	CGCCATCCCGG	GAAGATCTACGAGATCTACCTCACTCTGCACA	95084
QY	540	-----	-----	-----	544
Db	95083	AGCCAGAAAGCTG	AGGTGTGTGCC	CTGTGTGGGGGGGGGGGAGGCA	95024

QY 544 ----- 544  
Db 95023 CGTCTGTCTCTCTATCGCTGGAAACCAACCAATTCCTCCCGCCAGCACTGCTGTACCTG 94964  
QY 544 ----- 544  
Db 94963 TTCTGACCTGGCCCGAGGGAGCACAGGGATGAGGGCACTTAAGCTCACCTGATTGC 94904  
QY 544 ----- 544  
Db 94903 AACTTCCGTCCTACGGTGTCAACCAACCCCACTGCAGAGTGTGTCTCTAAACCCCAACGGT 94844  
QY 545 ----- ArgLeuP 547  
Db 94843 CTCTAGGACTGGCCACTCAACGCTTTCCCTCTCCACCCATGTTTCCCACTTGAGTTGC 94784  
QY 547 roLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValL 567  
Db 94783 CCTAGCTGGCTGTCAAGACCTGTCTGAGTCTCTATCGTTAGCTGTGGGCCCCAGAGTCC 94724  
QY 567 euLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerT 587  
Db 94723 TGCTCACCCGCCAGTCATCTTGCATGGACCACTGGGGGAGCCCACTGCCAGACACT 94664  
QY 587 rpSerLeuArgLeuLysLysGlnSerCysGlyGlySerTirpGlu----- 601  
Db 94663 GGAGCCTCGGCCTCAAAAAGCAGTCTCTGTAGGGCAGCTGGGAGGTGAGTTGGGAAGACT 94604  
QY 601 ----- 601  
Db 94603 ACCTGAGGCCACCAACCAATCACTGCCCCCACTGGGCATGGCCTCCCATTTCTTTG 94544  
QY 601 ----- 601  
Db 94543 GCAGCTATGTGGTCTCGAGTCTCCCAAACTGAGCTGAGGTGGCCCGAGGCTCTAAGG 94484  
QY 602 ----- GlnAspValLeuHisLeu 607  
Db 94483 ACTCCTAAGTTGGTTAGGCACGTACCACCAAGGCCCGCCACAGAGAGTGTGCACCTT 94424  
QY 608 GlyGluGluAlaProSerHisLeuTyTyCysGlnLeuGluAlaSerAlaCysTyVal 627  
Db 94423 GGTGAGAGTGGCCCTCTCATCTCTACTCTGAGGTGGAGGCGGGGCTGTATGTC 94364  
QY 628 PheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAla 647  
Db 94363 TTCACCGAGCAGTAGGCCGCTTTGCCCTGGTGAGAGGCCCTTCACGCTGGCTGCCACC 94304  
QY 648 LysArgLeuLysLeuLeuPheAlaProValAlaCysThrSerLeuGluTyAsnIle 667  
Db 94303 AAGCGCTCAGGCTCTCTTCTTTGCCCCCTGTGGCCTGTAGTCCCTCGAGTACACATC 94244  
QY 668 AtqValTyrcysLeuHisAspThrHisAspAlaLeuLys----- 680  
Db 94243 CGAGTGTACTGCTGCACGACACCCACGATGCTCTCAA - GGTACTGCGCGGCACAGCCCC 94185  
QY 680 ----- 680  
Db 94184 ATTGGCGGACCCACAGCCCTTGGCATCCAACTGCCTTGTCTTGGGGGTGACAGGAGGG 94125  
QY 681 ----- GluValValGlnLeu 685  
Db 94124 TTTGCTCTACCGTCGGGCCCTTGTGACACCCCGGGCAGCTGCAGGAGGTGGTGCAGCTG 94065  
QY 686 GluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705  
Db 94064 GAGAAGCAGCTGGGTGGACAGCTGATCCAGAGGCCCGTGTCTGTCTGCACTTCAAAAGACAGT 94005  
QY 706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTyPheLysLeu 725  
Db 94004 TACCACAACCTAGTCTGTCTCATCCAGAGTGCCTGCTGTTGGGAAGCAAGCTC 93945

QY 726 LeuValSerTyrGln----- 730  
Db 93944 CTTGTCACTACCA - GGTGAGACCTGTAGCCCAACGACAGAGCGCTCCAGGGCTGCTGG 93886  
QY 731 ----- GluIleProPheTyHis 736  
Db 93885 CGCTGGGCACGTTGACGCTCTCTTTTCCCTCTCCATAGGAGATCCCTTTTACCAC 93826  
QY 737 IleTirpAsnGlyThrGlnArgTyrIleuHisCysThrPheThrLeuGluArgValSerPro 756  
Db 93825 ATCTGGAATGCACTCAGCAGTATCTGCACTGCACCTTCCACCTGAGGCGCTCAATGCC 93766  
QY 757 SerThrSerAspLeuAlaCysLysLeuTirpValTirpGlnValGluGlyAspGlyGlnSer 776  
Db 93765 AGCACCGACCTGGCTGCAAGGTGGGTGTGGCAGTGGAGGTGGAGGAGATGGACAGAGC 93706  
QY 777 PheSerIleAsnPheAsnIleThrLys----- 785  
Db 93705 TTCAACATCAACTTTAATCATCAAGTGGCGGCACAAGATGGTGTGCTCCTAGTACCAG 93646  
QY 785 ----- 785  
Db 93645 CCATCCAGGAGGCTCCACCTCTCGCGGTAGTGCACCACCAAGGCGCTTAAGTTGA 93586  
QY 785 ----- 785  
Db 93585 GCTGCCCTTCTAGCTTTCTTCAAGTTTCTGGGACATAGGTGTAGGTGTATATAGGA 93526  
QY 785 ----- 785  
Db 93525 GCGGAGGAGGTAGAGTTGTCTGGGGCTCTGGAGGCGCTTCCCAAGAGGAGCTTTGC 93466  
QY 785 ----- 785  
Db 93465 ACCGGCCCCAAAGATAGGAGGAGGCGGAACTCTGACAGACTGCAGGTTGGGAGAAG 93406  
QY 785 ----- 785  
Db 93405 CATTTGGGCGAGGCGCACTGAGCTGAGCATAGCCATGATGTCAGCAGAGCAATAG 93346  
QY 785 ----- 785  
Db 93345 TGTGGGAGGAGGAGGTGTCTCCGTCACCAACTGTGGCCAGGTGACTTCATCTCTGAG 93286  
QY 785 ----- 785  
Db 93285 ATCTAAAGAGTAAAGCCCTGGTGTACAGGGAGAGGATGCCCTGTGCCCCCATTTGCTG 93226  
QY 785 ----- 785  
Db 93225 GGGTTATCTGGAAACTCCCTGCGGAGGCGGAGGGGTCCAGAGTTCTTAAGTGTACCTG 93166  
QY 786 ----- AspThrArgPheAlaGluLeuAlaLeuGluSerG 798  
Db 93165 TTGCAAAACATCCCATGATACAGACACAGAGTTGTCTGAAATGCTGGCTCTGAGAGTGA 93106  
QY 798 uAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgG 818  
Db 93105 AGGGGGGTCCCAAGCCCTGGTGGGCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGCA 93046  
QY 818 nlyValIleSerSerLeuAspProProCysArgArgGlyAlaAspTirpArgThrLeuAl 838  
Db 93045 AAGATCATTTACGACCTGGACCCACCTGCAAGCGGGGCGGCGGAGTGGCGAACTCTAGC 92986  
QY 838 aGlnLysLeuHisLeuAspSer----- 845  
Db 92985 CCAGAAACTTCACTTGGACAG - GTGGGTGGAGCTAGACAGAGGAGGCTGGTTGGGCAAG 92927  
QY 845 ----- 845  
Db 92926 CGAGGTAGACAGAAAGGAGGCGCTGTAGGCTCTGTGCCACAGGCTCTAGTGGTATTTCA 92867  
QY 846 ----- HisLeuSe 848

```
Db 92866 GGCAGCCTGGAGTGGAGCCAGACTACTGTACAGACTGCCCTCTTCTTTGCAGCCATCTTAG 92807
Qy 848 rPhePheAlaSerIysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHi 868
Db 92806 CTTCTTTGGCCTCCAGCCCAAGCCCTACAGCCNTGATCTCAACCTATGGAGGGCGGCA 92747
Qy 868 sPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAs 888
Db 92746 CTTCCCCAACGGCAACCTCGGCCAGCTGGCCGAGCTGTGGCCGACTGGGCCAGCCAGA 92687
Qy 888 pAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 92686 TGCTGGCCTCTTCACCGCTGTACAGAGGCCGAGTGC 92653
```

Search completed: October 6, 2004, 21:08:28  
Job time : 9114 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 18:17:48 ; Search time 100 Seconds  
(without alignments)  
2892.975 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGIWPALLGIVLAAML.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4787	100.0	899	11	US-09-970-944-2
2	4698.5	98.2	898	12	US-09-918-779-2
3	4698.5	98.2	898	16	US-10-624-932-2
4	4545.5	95.0	898	10	US-09-933-261-5
5	4545.5	95.0	898	11	US-09-970-944-13
6	4545.5	95.0	898	14	US-10-256-702-5
7	4545.5	95.0	898	14	US-10-240-154-16
8	4379.5	91.5	842	16	US-10-311-623-1
9	2819	58.9	544	11	US-09-970-944-14
10	2789.5	58.3	557	10	US-09-933-261-6
11	2705.5	56.5	931	11	US-10-970-944-5
12	2705.5	56.5	931	12	US-10-087-684-35
13	2705.5	56.5	931	12	US-09-972-211-121
14	2705.5	56.5	931	12	US-10-037-417-117
15	2705.5	56.5	931	12	US-10-037-417-117

16	2705.5	56.5	931	12	US-10-096-625-121
17	2705.5	56.5	1010	12	US-10-218-779-35
18	2680.5	56.0	931	11	US-09-970-944-16
19	2680.5	56.0	931	12	US-09-972-211-125
20	2680.5	56.0	931	12	US-10-096-625-125
21	2677.5	55.9	931	12	US-10-037-417-120
22	2673.5	55.8	931	11	US-09-970-944-17
23	2673.5	55.8	931	12	US-10-087-684-36
24	2673.5	55.8	931	12	US-10-218-779-36
25	2673.5	55.8	931	12	US-09-972-211-122
26	2673.5	55.8	931	12	US-10-037-417-118
27	2673.5	55.8	931	12	US-10-037-417-119
28	2673.5	55.8	931	12	US-10-096-625-122
29	2497	52.2	945	12	US-10-087-684-34
30	2497	52.2	945	12	US-10-218-779-34
31	2497	52.2	945	12	US-09-972-211-124
32	2497	52.2	945	12	US-10-037-417-121
33	2497	52.2	945	12	US-10-096-625-124
34	2493	52.1	945	12	US-10-087-684-33
35	2493	52.1	945	12	US-10-218-779-33
36	2493	52.1	945	12	US-09-972-211-123
37	2493	52.1	945	12	US-10-096-625-123
38	2490.5	52.0	946	12	US-09-972-211-56
39	2490.5	52.0	946	12	US-10-096-625-56
40	2490	52.0	943	10	US-09-933-261-7
41	2490	52.0	943	14	US-10-256-702-7
42	2484	51.9	933	12	US-10-087-684-2
43	2484	51.9	933	12	US-10-087-684-4
44	2484	51.9	933	12	US-10-218-779-2
45	2484	51.9	933	12	US-10-218-779-4

ALIGNMENTS

RESULT 1

US-09-970-944-2  
; Sequence 2, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Heriman, John L  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shinkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1e1 Proteins and Nucleic Acids Encoding Same and Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR FILING DATE: 60/237,862  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 899  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-970-944-2

Query Match	100.0%	Score	4787	DB	11	Length	899
Best Local Similarity	100.0%	Pred	No. 0				
Matches	899	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MAVRPGIWPALLGIVLAAMLRGSGAQSSATVANPVGANPDLPLPFLVPEPDIYVKNKP	60				
Db	1	MAVRPGIWPALLGIVLAAMLRGSGAQSSATVANPVGANPDLPLPFLVPEPDIYVKNKP	60				
QY	61	VLLVCXKAVPATQIFKCNCEWVRQVDHVIERTDSSGSGPTMEVRINVSQQKVFGL	120				
Db	61	VLLVCXKAVPATQIFKCNCEWVRQVDHVIERTDSSGSGPTMEVRINVSQQKVFGL	120				
QY	121	EYWCQCVANSSGTTKSKAYIRIARLKNFQEPKAEVLSFOGIVLPCRPPEGTPPAE	180				
Db	121	EYWCQCVANSSGTTKSKAYIRIARLKNFQEPKAEVLSFOGIVLPCRPPEGTPPAE	180				

```

181 VEWLNEDLVDPDLDPNNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSAAVIVY 240
181 VEWLNEDLVDPDLDPNNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSAAVIVY 240
241 VNGWSWTWTSVCSASCGRWKRSCTNPAPLNGAFCEGONVHDRTVSSLLVSVDG 300
241 VNGWSWTWTSVCSASCGRWKRSCTNPAPLNGAFCEGONVHDRTVSSLLVSVDG 300
301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLDSVADSSILTSQFQVSIKPSKADNPHL 420
361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLDSVADSSILTSQFQVSIKPSKADNPHL 420
421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFVS 480
421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFVS 480
481 RLSTQNYFRSLPRGTSNNMTYGTFFNLGRLMIPNTGISLIPDAIPRGKIYEIYLTJHK 540
481 RLSTQNYFRSLPRGTSNNMTYGTFFNLGRLMIPNTGISLIPDAIPRGKIYEIYLTJHK 540
541 PEDVRLPLAGCQTLTLLSPVSCGPPGVLLTRPVILAMDHCCGEPSPDSLSRLKKQSCGSW 600
541 PEDVRLPLAGCQTLTLLSPVSCGPPGVLLTRPVILAMDHCCGEPSPDSLSRLKKQSCGSW 600
601 EQDVHLGEEAPSHLYYCOLEASCYVFTQGLRFALVGEALSVAARLKLILLFAPVAC 660
601 EQDVHLGEEAPSHLYYCOLEASCYVFTQGLRFALVGEALSVAARLKLILLFAPVAC 660
661 TSLEYNIRVYCLHDTDALKEVWOLEKQGLQTOEPRVHLHFQKDSYNNLRLSIHDPSSL 720
661 TSLEYNIRVYCLHDTDALKEVWOLEKQGLQTOEPRVHLHFQKDSYNNLRLSIHDPSSL 720
721 WKSLLYSYQBIPIFYHWNQGTQRYLHCTFTLERYSPSTSDIACKLWVQVGDGQSFSIN 780
721 WKSLLYSYQBIPIFYHWNQGTQRYLHCTFTLERYSPSTSDIACKLWVQVGDGQSFSIN 780
781 FNITKDTFAELLALAESEAGVPALVGSAPKIPFLIRQKIISLDPDPCRRGADWRTLAQK 840
781 FNITKDTFAELLALAESEAGVPALVGSAPKIPFLIRQKIISLDPDPCRRGADWRTLAQK 840
841 LHLDSHLSFFASKSPPTAMILNLEARHFPNGNLSQAAAAVAGLQPDAGLFTVSEAEBC 899
841 LHLDSHLSFFASKSPPTAMILNLEARHFPNGNLSQAAAAVAGLQPDAGLFTVSEAEBC 899

```

RESULT 2

```

US-09-918-779-2
; Sequence 2, Application US/09918779
; Publication No. US20030064369A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Albrook, John
; APPLICANT: Lepage, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David

```

```

; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/09/918,779
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,697
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/263,662
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/281,645
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: ERT
; ORGANISM: Homo sapiens
; US-09-918-779-2

```

Query Match 98.2%; Score 4698.5; DB 12; Length 898;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

```

Qy 1 MAVRPGIWPALLGIVLAALIRGSGAQSATVANVPVPGANPDLLPHFLVEPEDVTVKPK 60
Db 1 MAVRPGIWPALLGIVLAALIRGSGAQSATVANVPVPGANPDLLPHFLVEPEDVTVKPK 60
Qy 61 VLLVCKAVPATQIFPKCNGEVRQVDHVIERTDGSSEPTMEVRINVSQOQKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNGEVRQVDHVIERTDGSSEPTMEVRINVSQOQKVFGL 120
Qy 121 EYWCQCVAMSSSGTTKQKAYIRIARLRKNEQEPLEKEVSLQGIPLPCRPPEGIPPAE 180
Db 121 EYWCQCVAMSSSGTTKQKAYIRIARLRKNEQEPLEKEVSLQGIPLPCRPPEGIPPAE 180
Qy 181 VEWLNEDLVDPDLDPNNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSAAVIVY 240
Db 181 VEWLNEDLVDPDLDPNNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSAAVIVY 240
Qy 241 VNGWSWTWTSVCSASCGRWKRSCTNPAPLNGAFCEGONVHDRTVSSLLVSVDG 300
Db 241 VNGWSWTWTSVCSASCGRWKRSCTNPAPLNGAFCEGONVHDRTVSSLLVSVDG 300
Qy 301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
Db 301 SWPWSKWSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEDV 360
Qy 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLDSVADSSILTSQFQVSIKPSKADNPHL 420
Db 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLDSVADSSILTSQFQVSIKPSKADNPHL 420
Qy 421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFV 479
Db 421 LTIQPDLSLTITTYOGSLCPRODGPSPKFQLTNGHLLSPLGGGRHTLHSSPTSABEFV 479

```

QY 480 SRLSTONYFRSLPRGTSNMTYGTNFGLGRMLMINTGISTLLIPDAIPRGKIYIYLTILH 539  
Db 480 SRLSTONYFRSLPRGTSNMTYGTNFGLGRMLMINTGISTLLIPDAIPRGKIYIYLTILH 539  
QY 540 KPEDVRLPLACQOTLLSPIVSCGPGVULLTRPVILAMDHCGEPSPDSWSLRKQSCGFS 599  
Db 540 KPEDVRLPLACQOTLLSPIVSCGPGVULLTRPVILAMDHCGEPSPDSWSLRKQSCGFS 599  
QY 600 WEQDVLHLGEEAPSHLYYCOLEASACVYFTBOLGREALVGEALSVAARLKLILFAPVA 659  
Db 600 WE-DVLHLGEEAPSHLYYCOLEASACVYFTBOLGREALVGEALSVAARLKLILFAPVA 658  
QY 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGLIQEPVILHFKDSYHNLFSLHIDVPSS 719  
Db 659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGLIQEPVILHFKDSYHNLFSLHIDVPSS 718  
QY 720 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTFLERVSPSTDLACKLWVWQVGGQGSFSI 779  
Db 719 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTFLERVSPSTDLACKLWVWQVGGQGSFSI 778  
QY 780 NFNITKDTFRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 839  
Db 779 NFNITKDTFRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 838  
QY 840 KLHLDLSLSPFASKPSTAMILNLWEARHPNGLSOLAAVAGLQGFDAGLFTVSEAC 899  
Db 839 KLHLDLSLSPFASKPSTAMILNLWEARHPNGLSOLAAVAGLQGFDAGLFTVSEAC 898

## RESULT 3

US-10-624-932-2  
; Sequence 2, Application US/10624932  
; Publication No. US20040096877A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier, Raymond  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Grosse, William  
; APPLICANT: Alsobrook, John  
; APPLICANT: Lepley, Denise  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David  
; APPLICANT: Smithson, Glennda  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-074 US  
; CURRENT APPLICATION NUMBER: US/10/624,932  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: 09/918,779  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: 60/221,409  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 60/222,840  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,752  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,762  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,770  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,769  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/225,146  
; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/225,392  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/225,470  
; PRIOR FILING DATE: 2000-08-15  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 898  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-624-932-2

Query Match 98.2%; Score 4698.5; DB 16; Length 898;  
Best Local Similarity 98.7%; Pred. No. 0;

Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

QY 1 MAVRPGLPALLGLVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPEDVYIVKNKP 60  
Db 1 MAVRPGLPALLGLVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPEDVYIVKNKP 60  
QY 61 VLLVCKAVPATQIFFKCNQGEVVRQDVHVIERSSTGSSGEPTMEVRIINVSRQQVEKVFGL 120  
Db 61 VLLVCKAVPATQIFFKCNQGEVVRQDVHVIERSSTGSSGLPTMEVRIINVSRQQVEKVFGL 120  
QY 121 EYWCQCQVAMSSGTTKSOKAYIRIARLRKNFBOEPLAKVSLQGVILPCRPPEGIPPAE 180  
Db 121 EYWCQCQVAMSSGTTKSOKAYIRIARLRKNFBOEPLAKVSLQGVILPCRPPEGIPPAE 180  
QY 181 VEWLNRNEDLVDPNVIYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240  
Db 181 VEWLNRNEDLVDPNVIYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240  
QY 241 VNGGWSWTWSEVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHDRTVSSILVSDVG 300  
Db 241 VNGGWSWTWSEVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299  
QY 301 SWSFWSKWSACGLDCTHWSRECSDPAPRNGBECQGTDLDTNCTSDLCVHSASGEDV 360  
Db 300 SWSFWSKWSACGLDCTHWSRECSDPAPRNGBECQGTDLDTNCTSDLCVHSASGEDV 359  
QY 361 ALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDVADSSILTSFGFOPVSIKPSKADNPHL 420  
Db 360 ALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDVADSSILTSFGFOPVSIKPSKADNPHL 419  
QY 421 LTIQPDLS-TTTTYQSLCPROQDPSPKFQLTNGHLLSPILGGGRHTLHSSPTSEAEFV 479  
Db 420 LTIQPDLS-TTTTYQSLCPROQDPSPKFQLTNGHLLSPILGGGRHTLHSSPTSEAEFV 479  
QY 480 SRLSTONYFRSLPRGTSNMTYGTNFGLGRMLMINTGISTLLIPDAIPRGKIYIYLTILH 539  
Db 480 SRLSTONYFRSLPRGTSNMTYGTNFGLGRMLMINTGISTLLIPDAIPRGKIYIYLTILH 539  
QY 540 KPEDVRLPLAGCQOTLLSPIVSCGPGVULLTRPVILAMDHCGEPSPDSWSLRKQSCGFS 599  
Db 540 KPEDVRLPLAGCQOTLLSPIVSCGPGVULLTRPVILAMDHCGEPSPDSWSLRKQSCGFS 599  
QY 600 WEQDVLHLGEEAPSHLYYCOLEASACVYFTBOLGREALVGEALSVAARLKLILFAPVA 659  
Db 600 WE-DVLHLGEEAPSHLYYCOLEASACVYFTBOLGREALVGEALSVAARLKLILFAPVA 658  
QY 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGLIQEPVILHFKDSYHNLFSLHIDVPSS 719  
Db 659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGLIQEPVILHFKDSYHNLFSLHIDVPSS 718  
QY 720 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTFLERVSPSTDLACKLWVWQVGGQGSFSI 779  
Db 719 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTFLERVSPSTDLACKLWVWQVGGQGSFSI 778  
QY 780 NFNITKDTFRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 839  
Db 779 NFNITKDTFRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTLAQ 838

QY 840 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLQPDAGLFTVSEAC 899  
DB 839 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLQPDAGLFTVSEAC 898

RESULT 4  
US-09-933-261-5  
; Sequence 5, Application US/09933261  
; Publication No. US20030040046A1  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsey  
; Masu, Masayuki  
; Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/933,261  
; FILING DATE: 20-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/808,982  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 898 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: No. US20030040046A1 Relevant  
; TOPOLOGY: No. US20030040046A1 Relevant  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-933-261-5

Query Match 95.0%; Score 4545.5; DB 10; Length 898;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPLMPALLGIVLAALRGSGAQQSATVANPVGANPDLLPHFLVEPDVYVKNKP 60  
DB 1 MAVRPLMPVLLGIVLAALRGSGAQQSATVANPVGANPDLLPHFLVEPDVYVKNKP 60

QY 61 VLLVCKAVPATQIFKNGEVRQVDHVIERTSDGSGEPTMEVRINVSROQVEKVGLE 120  
DB 61 VLLVCKAVPATQIFKNGEVRQVDHVIERTSDGSGEPTMEVRINVSROQVEKVGLE 120

QY 121 EYWCOCVWSSGGTTKSKAVIRIARLRNKEQEPLAKEVSLQGIQVLCRPPGIPPAE 180  
DB 121 EYWCOCVWSSGGTTKSKAVIRIARLRNKEQEPLAKEVSLQGIQVLCRPPGIPPAE 180

QY 181 VEWLNRNEDVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240  
DB 181 VEWLNRNEDVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAIVY 240

QY 241 VNGGWTWTWSEVCSASCGRWQKRSCCTNPAPLNGAFCEQNVHRTVSSILVSDG 300  
DB 241 VNGGWTWTWSEVCSASCGRWQKRSCCTNPAPLNGAFCEQNV-QKTACATLCPVDG 299

QY 301 SWSPSKWSACGLDCTHWSRECSDPAPRNGBECQGTDLTRNCTSDLCVHSGASPEDV 360  
DB 300 SWSWSKWSACGLDCTHWSRECSDPAPRNGBECRGADLTNCTSDLCVHSGASPEDV 359

QY 361 ALYVGLIYAVAVCLVLLVLLVLYCRKKKGLDSDVADSSILTSQFOPVSTKPSKADNPHL 420  
DB 360 ALYIGLVAVAVCLVLLVLLVLYCRKKKGLDSDVADSSILTSQFOPVSTKPSKADNPHL 419

QY 421 LTIQPDLS-TTTTYQGSCLCPRODGPSPKFOLITNGHLLSPGLGGGRHTLHSSPTSEAEFV 479  
DB 420 LTIQPDLSITTTTYQGSCLCSRDGPSPKFOLSGRHTLHSSPTSEAEFV 479

QY 480 SRLSTQNYFRSLPRGTSNMTYGTNFGGLGRIMIPNTGISLLIPPDALIPRKIYEIYTLH 539  
DB 480 SRLSTQNYFRSLPRGTSNMTYGTNFGGLGRIMIPNTGISLLIPPDALIPRKIYEIYTLH 539

QY 540 KPEDVRLPAGCOTLLSPYVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLLKQSCGS 599  
DB 540 KPEDVRLPAGCOTLLSPVYVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLLKQSCGS 599

QY 600 WEODVLHGBEAPSHLYVCOLEASACVYFTEQLGRFALVGEALSVAALKLILFAPVA 659  
DB 600 WE-DVLHGBEAPSHLYVCOLEAGACVYFTEQLGRFALVGEALSVAALKLILFAPVA 658

QY 660 CTSLEYNIRVYCLHDTDALKEVVVQLEKQGGQIQBPVRLVFKDSVHNLRLSHDVPS 719  
DB 659 CTSLEYNIRVYCLHDTDALKEVVVQLEKQGGQIQBPVRLVFKDSVHNLRLSHDVPS 718

QY 720 LWSKLLVSYQETPFYHINWGTQRYLHCTETLERSVSTSDLACKLVWVQVGGQGSFI 779  
DB 719 LWSKLLVSYQETPFYHINWGTQRYLHCTETLERSVSTSDLACKLVWVQVGGQGSFI 778

QY 780 NFNTKDTFAELIALESEAAGVPALVGPSAFKIPFLIRQKIISLDPCCRGADWRTLAQ 839  
DB 779 NFNTKDTFAELIALESEAAGVPALVGPSAFKIPFLIRQKIISLDPCCRGADWRTLAQ 838

QY 840 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLQPDAGLFTVSEAC 899  
DB 839 KLHLDHLSFFASKPSPTAMILNWEARHPNGLSOLAAAVAGLQPDAGLFTVSEAC 898

RESULT 5  
US-09-970-944-13  
; Sequence 13, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Herrman, John L  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and  
; TITLE OF INVENTION: Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/237,862  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 898  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-970-944-13

Query Match 95.0%; Score 4545.5; DB 11; Length 898;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPLMPALLGIVLAALRGSGAQQSATVANPVGANPDLLPHFLVEPDVYVKNKP 60



```
Db 1 MAVRPGLMPVLLGIVLAAMLRGSAQASATVANPVGANPDLLPHFLVEPEDYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERSSTGSSGSEPTMEVRINVSROQVEKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERSSTGSSGSEPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCQCVAMSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIVLPCRPPGIPPAE 180
Db 121 EYWCQCVAMSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIVLPCRPPGIPPAE 180
QY 181 VEWLNRDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
Db 181 VEWLNRDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGAFCEGQNV-QKTACATLCPVDG 300
Db 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGAFCEGQNV-QKTACATLCPVDG 300
QY 301 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPEV 360
Db 300 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPEV 360
QY 361 ALYVGLTAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
Db 360 ALYVGLTAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
QY 421 LTIQPDLS-TTTTYQGSICPRQDPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
Db 420 LTIQPDLS-TTTTYQGSICPRQDPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSMYTGTFNFGRLMIPNTGISLIPDPAIPRGKIYIYLT 539
Db 480 SRLSTQNYFRSLPRGTSMYTGTFNFGRLMIPNTGISLIPDPAIPRGKIYIYLT 539
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKQSCGS 599
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKQSCGS 599
QY 600 WEODVHLGEPASHLYCOLEASACVFTQGLRFPALVGEALSVAAKBLKLLFAPVA 659
Db 600 WE-DVLHGESFSLHYCOLEAGACVFTQGLRFPALVGEALSVAAKBLKLLFAPVA 658
QY 660 CTSLEYNIRYCLDTHDALKEVVQLEKQGGQIIOEPVRLHFKDSVHNLRLSHDVPSS 719
Db 659 CTSLEYNIRYCLDTHDALKEVVQLEKQGGQIIOEPVRLHFKDSVHNLRLSHDVPSS 718
QY 720 LWSKLLVSYQEIPIFYHWMNGTQYLHCTFTLERINASTDLACKVWVQVGDGQSFNI 779
Db 719 LWSKLLVSYQEIPIFYHWMNGTQYLHCTFTLERINASTDLACKVWVQVGDGQSFNI 778
QY 780 NFENITKDTFAELALESEAGVPALVCPSAFKIPFLIRQKIISLDPCCRGADWRTLAQ 839
Db 779 NFENITKDTFAELALESEAGVPALVCPSAFKIPFLIRQKIISLDPCCRGADWRTLAQ 838
QY 840 KLHLDLHLSFFASKPSPTAMILNLWEARHPNGLSOLAAVAGLQPDAGLFTVSEAE 899
Db 839 KLHLDLHLSFFASKPSPTAMILNLWEARHPNGLSOLAAVAGLQPDAGLFTVSEAE 898

RESULT 6
US-10-256-702-5
; Sequence 5, Application US/10256702
; Publication No. US20030059859A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
```

```
;
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059859A1 Relevant
; TOPOLOGY: No. US20030059859A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
;
US-10-256-702-5
```

```
Query Match 95.0%; Score 4545.5; DB 14; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
```

```
QY 1 MAVRPGLMPALLGIVLAAMLRGSAQASATVANPVGANPDLLPHFLVEPEDYIVKNKP 60
Db 1 MAVRPGLMPVLLGIVLAAMLRGSAQASATVANPVGANPDLLPHFLVEPEDYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERSSTGSSGSEPTMEVRINVSROQVEKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNQWVRQVDHVIERSSTGSSGSEPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCQCVAMSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIVLPCRPPGIPPAE 180
Db 121 EYWCQCVAMSSSGTTKSKAYIRIARLNKNEQEPLEKEVLEQIVLPCRPPGIPPAE 180
QY 181 VEWLNRDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
Db 181 VEWLNRDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGAFCEGQNVHDTVSSLLVSDG 300
Db 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPNGGAFCEGQNV-QKTACATLCPVDG 299
QY 301 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPEV 360
Db 300 SWSPWSKWSACGLDCTHWRSECSDPAPNGGECQGTDLTRNCTSDLCVHSASGPEV 359
QY 361 ALYVGLTAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420
Db 360 ALYVGLTAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTTTYQGSICPRQDPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
Db 420 LTIQPDLS-TTTTYQGSICPRQDPSPKFQLTNGHLLSPGGRHTLHSSPTSEAEFV 479
```

```

QY 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
DB 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
QY 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDWSRLKKQSCGS 599
DB 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDWSRLKKQSCGS 599
QY 600 WEODVHLGGEAPSHLYYCOLEASACVFTFQGRFALVGEALSVAALKLLPAPVA 659
DB 600 WE-DVLHLGGEAPSHLYYCOLEASACVFTFQGRFALVGEALSVAALKLLPAPVA 658
QY 660 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 719
DB 660 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 718
QY 720 LWSKLLVSYQEIIPFYHWMGTQYLLHCTFTLERVSPSTSDACKLWVWQVEGDGQSFNI 779
DB 720 LWSKLLVSYQEIIPFYHWMGTQYLLHCTFTLERVSPSTSDACKLWVWQVEGDGQSFNI 778
QY 780 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 839
DB 780 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 838
QY 840 KLHLDHSLSFASKPSPMTAMLNWEARHPNGLSQLAAVAGLGPDPAGLFTVSEAE 899
DB 840 KLHLDHSLSFASKPSPMTAMLNWEARHPNGLSQLAAVAGLGPDPAGLFTVSEAE 898

RESULT 7
US-10-240-154-16
; Sequence 16, Application US/10240154
; Publication No. US2003017541A1
; GENERAL INFORMATION:
; APPLICANT: Cochran et al.
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
; FILE REFERENCE: CFM-P01-006
; CURRENT APPLICATION NUMBER: US/10/240,154
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/GB01/01486
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-240-154-16

Query Match 95.0%; Score 4545.5; DB 14; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRGLWPAALGLVLAALWLGSGAQSSATVANPVPANPOLLPHFLVEPDEVIVKNKP 60
DB 1 MAVRGLWPAALGLVLAALWLGSGAQSSATVANPVPANPOLLPHFLVEPDEVIVKNKP 60
QY 61 VLLVCXAVPATOIFPKCNGEVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
DB 61 VLLVCXAVPATOIFPKCNGEVRQVDHVIERTDSSGSEPTMEVRINVSROQVEKVFGL 120
QY 121 EYWCQCVANSSGTTKSKAVIRIARLKNPEOEPLAKEVSLQGIPLPCRPPEGIPPAE 180
DB 121 EYWCQCVANSSGTTKSKAVIRIARLKNPEOEPLAKEVSLQGIPLPCRPPEGIPPAE 180
QY 181 VEVLNEDLVPSLDPNVIITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
DB 181 VEVLNEDLVPSLDPNVIITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY 241 VNGGSHSTWTEWSCASGRGQKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 300
DB 241 VNGGSHSTWTEWSCASGRGQKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 300

```

```

QY 301 SMSWPKSWACGLDCTHWSRECSDPAPRNGGECQGTDLDTNCTSDLCVHSASPEDV 360
DB 300 SMSWPKSWACGLDCTHWSRECSDPAPRNGGECQGTDLDTNCTSDLCVHSASPEDV 359
QY 361 ALYVGLIAVAVCVLVLVLLVLYCKKGLDSDVADSSILTSFGQPVSIKPSKADNPHL 420
DB 360 ALYVGLIAVAVCVLVLVLLVLYCKKGLDSDVADSSILTSFGQPVSIKPSKADNPHL 419
QY 421 LTTQPDLS-TTTTYQSLCPRQDPSPKLTNTGNHLLSLPLGGGRHTLHSSPSTSEAEFV 479
DB 420 LTTQPDLS-TTTTYQSLCPRQDPSPKLTNTGNHLLSLPLGGGRHTLHSSPSTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
DB 480 SRLSTQNYFRSLPRGTSNMTYCTFNFLGGLMIPNTGISLILPPDAIPRCKIYEIYTLH 539
QY 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDWSRLKKQSCGS 599
DB 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPSDWSRLKKQSCGS 599
QY 600 WEODVHLGGEAPSHLYYCOLEASACVFTFQGRFALVGEALSVAALKLLPAPVA 659
DB 600 WE-DVLHLGGEAPSHLYYCOLEASACVFTFQGRFALVGEALSVAALKLLPAPVA 658
QY 660 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 719
DB 660 CTSLEYNIRVYCLDTHDALKEVVQLEKQGGQIQQEPRVLHFKDSVHNLRLSLHDPVSS 718
QY 720 LWSKLLVSYQEIIPFYHWMGTQYLLHCTFTLERVSPSTSDACKLWVWQVEGDGQSFNI 779
DB 720 LWSKLLVSYQEIIPFYHWMGTQYLLHCTFTLERVSPSTSDACKLWVWQVEGDGQSFNI 778
QY 780 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 839
DB 779 NFNITKDTREAFELALAESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRGADWRTIAQ 838
QY 840 KLHLDHSLSFASKPSPMTAMLNWEARHPNGLSQLAAVAGLGPDPAGLFTVSEAE 899
DB 839 KLHLDHSLSFASKPSPMTAMLNWEARHPNGLSQLAAVAGLGPDPAGLFTVSEAE 898

RESULT 8
US-10-311-623-1
; Sequence 1, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Dannel B.
; APPLICANT: TANG, Y. Tom; LAU, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Aneena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANTANWALA, Madhusudan M.
; APPLICANT: DUGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program

```

; SEQ ID NO 1  
; LENGTH: 842  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CD1  
US-10-311-623-1

Query Match  
Best Local Similarity 91.5%; Score 4379.5; DB 16; Length 842;  
Matches 838; Conservative 1; Mismatches 2; Indels 59; Gaps 3;  
QY 1 MAVRPGLWALLGIVLAALRGSAQQAQATVANPVGANPOLLPHFLVEPEDVIVKKNP 60  
DB 1 MAVRPGLWALLGIVLAALRGSAQQAQATVANPVGANPOLLPHFLVEPEDVIVKKNP 60  
QY 61 VLLVCKAVPATQIFPKNGEWRQVDHVIERTSGSGSEPTMEVRIIVSRQQVEKVFGL 120  
DB 61 VLLVCKAVPATQIFPKNGEWRQVDHVIERTSGSGSEPTMEVRIIVSRQQVEKVFGL 120  
QY 121 EYWCQVAMSSGTTKQKAVIRIARLKNPEQBLAKEVLSLEQIVLPCRPPGIPPAE 180  
DB 121 EYWCQVAMSSGTTKQKAVIRIARLKNPEQBLAKEVLSLEQIVLPCRPPGIPPAE 180  
QY 181 VEMLRNEDLVDPSPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240  
DB 181 VEMLRNEDLVDPSPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240  
QY 241 VNGGWSTWTSVCSACGRGWKRSCTNPAPLNGAFCEGQNVHDTVSSLLVSDG 300  
DB 241 -----VDG 243  
QY 301 SWSPWKSACGLDCTHWRSECDPAPRNGEECQGTDLDRNCTSDLCVHSASGPEV 360  
DB 244 SWSPWKSACGLDCTHWRSECDPAPRNGEECQGTDLDRNCTSDLCVHTASGPEV 303  
QY 361 ALVYGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNPHL 420  
DB 304 ALVYGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNPHL 363  
QY 421 LTTPQDLS -TTTTYQGSICPRQDPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAEFF 479  
DB 364 LTTPQDLS -TTTTYQGSICPRQDPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAEFF 423  
QY 480 SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGILIPDAIPRGKIYEIYLT 539  
DB 424 SRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGILIPDAIPRGKIYEIYLT 483  
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 599  
DB 484 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 543  
QY 600 WEQDVHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAALKLKLFPAP 659  
DB 544 WE -DVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAALKLKLFPAP 602  
QY 660 CTSLEYNIRVYCLDTHDALKEVVOLEKQGGQIIQEPVRLHFKDSYHNLRSLIHDVPS 719  
DB 603 CTSLEYNIRVYCLDTHDALKEVVOLEKQGGQIIQEPVRLHFKDSYHNLRSLIHDVPS 662  
QY 720 LWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVGGQGSFI 779  
DB 663 LWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVGGQGSFI 722  
QY 780 NFNITKDTFAELLALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839  
DB 723 NFNITKDTFAELLALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 782  
QY 840 KLHLDLSHLSFFASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQPDAGLFTVSEAC 899  
DB 783 KLHLDLSHLSFFASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQPDAGLFTVSEAC 842

## RESULT 9

US-09-970-944-14  
; Sequence 14, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Herrman, John L  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shmukets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and  
; FILE OF INVENTION: Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/237,862  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-970-944-14

Query Match 58.9%; Score 2819; DB 11; Length 544;  
Best Local Similarity 99.6%; Pred. No. 4.2e-238;  
Matches 540; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 359 DVALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNP 418  
DB 4 DVALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSQFPVSIKPSKADNP 63  
QY 419 HLLTIQDLS -TTTTYQGSICPRQDPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAE 477  
DB 64 HLLTIQDLS -TTTTYQGSICPRQDPSPKFQLTNGHLLSPLGGGRHTLHSSPTSEAE 123  
QY 478 FVSRSLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGILIPDAIPRGKIYEIYLT 537  
DB 124 FVSRSLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMPNTGILIPDAIPRGKIYEIYLT 183  
QY 538 LHPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSC 597  
DB 184 LHPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSC 243  
QY 598 GSWEQDVHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAALKLKLFPAP 657  
DB 244 GSWE -DVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAALKLKLFPAP 302  
QY 658 VACTSLEYNIRVYCLDTHDALKEVVOLEKQGGQIIQEPVRLHFKDSYHNLRSLIHDV 717  
DB 303 VACTSLEYNIRVYCLDTHDALKEVVOLEKQGGQIIQEPVRLHFKDSYHNLRSLIHDV 362  
QY 718 SSLWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVGGQGSF 777  
DB 363 SSLWSKLLVSVQEIIPFYHWNQYRLHCTFTLERVSPSTSDACKLWVWQVGGQGSF 422  
QY 778 SINFNITKDTFAELLALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRT 837  
DB 423 SINFNITKDTFAELLALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRT 482  
QY 838 AQKHLDSHLSFFASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQPDAGLFTVSE 897  
DB 483 AQKHLDSHLSFFASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQPDAGLFTVSE 542  
QY 898 EC 899  
DB 543 EC 544

## RESULT 10

US-09-933-261-6  
; Sequence 6, Application US/09933261  
; Publication No. US20030040046A1

```
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David
Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/933,261
FILING DATE: 20-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/808,982
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030040046A1 Relevant
TOPOLOGY: No. US20030040046A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-933-261-6
Query Match 58.3%; Score 2789.5; DB 10; Length 557;
Best Local Similarity 96.4%; Pred. No. 1.7e-235;
Matches 538; Conservative 2; Mismatches 15; Indels 3; Gaps 3;
QY 344 NCTSDLVHSAGSPEDVALYVGLIAVAVCLVLLVLLVLYVCRKKEGLDSDVADSSILTS 403
Db 1 NCTSDLVHSAGSPEDVALYVGLIAVAVCLVLLVLLVLYVCRKKEGLDSDVADSSILTS 60
QY 404 GFQPVSIKPSKADNPILLTIQPDLS-TTTTYQSLCPQDGPSPKFQLTNGHLLSPLGGG 462
Db 61 GFQPVSIKPSKADNPILLTIQPDLS-TTTTYQSLCPQDGPSPKFQLTNGHLLSPLGGG 120
QY 463 RHTLHSSPTSEAEFVSRSLTQNYFRSLPRGTSNMTYGTNFTGLGRMLPNTGSLIIP 522
Db 121 RHTLHSSPTSEAEFVSRSLTQNYFRSLPRGTSNMTYGTNFTGLGRMLPNTGSLIIP 180
QY 523 PDATPRGKIYIYLTTHKPEDVRLPLAGCQTLLSPVSCGPPGVLLTPFVILAMDHCGEP 582
Db 181 PDATPRGKIYIYLTTHKPEDVRLPLAGCQTLLSPVSCGPPGVLLTPFVILAMDHCGEP 240
QY 583 SPDSWSLRKQSCGSGWEQDVLHIGEPAPSHLYYQLEASACYVFTQLGRFALVGEAL 642
Db 241 SPDSWSLRKQSCGSGWE-DVLHIGEPAPSHLYYQLEASACYVFTQLGRFALVGEAL 299
QY 643 SVAAAKRLKLLFAPVACTSYENIRVYVCLHDTALKEVVOLEKQGLIQEPRVLHF 702
Db 300 SVAAAKRLKLLFAPVACTSYENIRVYVCLHDTALKEVVOLEKQGLIQEPRVLHF 359
QY 703 KDSYHNLRSLIHDVPSSLWKSLLVSYQEIFYHWNQYVHNGTQRYLHCTFTLERSVSPSTDLA 762
Db 1 KDSYHNLRSLIHDVPSSLWKSLLVSYQEIFYHWNQYVHNGTQRYLHCTFTLERSVSPSTDLA 419
QY 763 CKLWVQVVEGDGQSFNINFTKDTFAELLALLESAGVPAIVGSPSAFKIPFLIRQKLI 822
Db 420 CKLWVQVVEGDGQSFNINFTKDTFAELLALLESAGVPAIVGSPSAFKIPFLIRQKLI 479
QY 823 SLDPPCRGGADWRTLAQKHLDSHLSFFASKPSPTAMILNLWEARHPNGLSOLAAVA 882
Db 480 SLDPPCRGGADWRTLAQKHLDSHLSFFASKPSPTAMILNLWEARHPNGLSOLAAVA 539
QY 883 GLGQPDAGLFT-VSEAE 899
Db 540 GTXPAGRWLLSQCSAEAC 557
RESULT 11
US-10-256-702-6
Sequence 6, Application US/10256702
Publication No. US20030059859A1
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
Leonardo, E. David
Hink, Lindsay
Masu, Masayuki
Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/256,702
FILING DATE: 27-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/933,261
FILING DATE: 20-Aug-2001
APPLICATION NUMBER: 08/808,982
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 557 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030059859A1 Relevant
TOPOLOGY: No. US20030059859A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-256-702-6
Query Match 58.3%; Score 2789.5; DB 14; Length 557;
Best Local Similarity 96.4%; Pred. No. 1.7e-235;
Matches 538; Conservative 2; Mismatches 15; Indels 3; Gaps 3;
QY 344 NCTSDLVHSAGSPEDVALYVGLIAVAVCLVLLVLLVLYVCRKKEGLDSDVADSSILTS 403
Db 1 NCTSDLVHSAGSPEDVALYVGLIAVAVCLVLLVLLVLYVCRKKEGLDSDVADSSILTS 60
```

QY 404 GFQVSIKPSKADNPHLLTIQPDLS-FTTYYQSLCPQDGPSPKPOLTNHLLSPGLGG 462  
DB 61 GFQVSIKPSKADNPHLLTIQPDLSFTTYYQSLCPQDGPSPKPOLTNHLLSPGLGG 120  
QY 463 RHTLHSSPTSEAEFEFVSRSLSTQNYFSLPRGTSNMTYGTNFGLGRLMIPNTGISLLIP 522  
DB 121 RHTLHSSPTSEAEFEFVSRSLSTQNYFSLPRGTSNMTYGTNFGLGRLMIPNTGISLLIP 180  
QY 523 PDAIPRGKIYEIYLTLLHKPDVRLPLAGCQTLSPVSCGPPGVLLTRPVLAMDHGCEP 582  
DB 181 PDAIPRGKIYEIYLTLLHKPDVRLPLAGCQTLSPVSCGPPGVLLTRPVLAMDHGCEP 240  
QY 583 SPDSWSLRLKKQSCGSEWODVLHGBEAPSHLYYCOLEASACVVFTEQLGRFALVGEAL 642  
DB 241 SPDSWSLRLKKQSCGSEW- DVLHGBEAPSHLYYCOLEASACVVFTEQLGRFALVGEAL 299  
QY 643 SVAARLKLALLFAPVACTSLEYNIRVYCLDTHDALKEVVQLEKQLGGQLIQEPRVILH 702  
DB 300 SVAARLKLALLFAPVACTSLEYNIRVYCLDTHDALKEVVQLEKQLGGQLIQEPRVILH 359  
QY 703 KDSVHNLRLSIHDVPSLSLWKSLLVSYQETPFVHIWNGTORYLHCTFTLERSVSPSTDIA 762  
DB 360 XDSTHNLXLSHDVPSLSLWKSLLVSYQETPFVHIWNGTORYLHCTFTLERSVSPSTDIA 419  
QY 763 CKLWVQVEGDGQSFNFINITKDTFAELLALASEAGVPALVGPFAFKIPFLIROKILIS 822  
DB 420 CKLWVQVEGDGQSFNFINITKDTFAELLALASEAGVPALVGPFAFKIPFLIROKILIS 479  
QY 823 SLDPCCRRGADWRTLAKGLHDLSHLGFASPKPSPTAMILNLWEARHPNGLSQAALAAVA 882  
DB 480 SLDPCCRRGADWRTLAKGLHDLSHLGFASPKPSPTAMILNLWEARHPNGLSQAALAAVA 539  
QY 883 GLGQPDAGLFT-VSEAE 899  
DB 540 GTXPAGRWLLSQSEAE 557

## RESULT 12

US-09-970-944-15  
; Sequence 15, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Heriman, John L  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shinkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an  
; TITLE OF INVENTION: Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/237,862  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 931  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-970-944-15

Query Match 56.5%; Score 2705.5; DB 11; Length 931;  
Best Local Similarity 56.4%; Pred. No. 9e-228;  
Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;

QY 9 PALLGIVLAWLGRGSAQQA---TVANVPVGPANPDLLPHFLVEPEDVIVTKNKPVLVLC 65  
DB 26 PAL--ALLSASGTGSAQDDPFPHLPETFPSPDPPEPLPHFLIEPEAVIVKNKPNVLYC 83  
QY 66 KAVPATQIFPKCNKGVWRQVDHVIERSSTGSGSEPTMEVRINVSROQVKVFGLEYWCQ 125  
DB 84 KASPATQIFPKCNKSEWHQKDHVDRVDETSGLIVREVSIBISROQVBEELFGPDYWCQ 143

QY 126 CVAWSSGTTKSKAYIARLARKNKEOEPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR 185  
DB 144 CVAWSSAGITTSKAYVRIAYLRKTFEQEPGKEVSLQEVLLQCRPEGIPPAEVEWLR 203  
QY 186 NEDLVDPSPDPNYYITREHSLVVRQARLADTANYTCVAKNIVARRRRSASAAVIVYVNGW 245  
DB 204 NEDIIDPAEDRNFYIIDHNLIIKQARLSTANYTCVAKNIVAKKSTTATVIVYVNGW 263  
QY 246 STWTEWSVCASGCRGQKRSRSTNPAPLNGAFCEGQVNDRTVSSLLVSDGSPW 305  
DB 264 STWTEWSVCNRCRGYQKTRTCTNPAPLNGAFCEGQSV-QKIACTTLCFVDGRWTSW 322  
QY 306 SKWSACGLDCTHWRSECDPAPRNGECCQGTDLTRNCTSDLCVHASGSPEDVALYVG 365  
DB 323 SKWSTGTECTHWRRECTAPAPKNGKCDGLVLSQKNCTDGLCQMAAPSDDDVALYVG 382  
QY 366 L-TAVAVCLVLLLVLLVYCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHLLTIQ 424  
DB 383 IVIAVTVCIAITVVVALFVYKRNHRDFESDIIDSSALNGGFQVNIKAARQD---LLAVP 439  
QY 425 PDL-STYTTTYQSLCPQDGPSPKPOLTNHLLSPGLGGRLTHHSS----PTSEAEFV 479  
DB 440 PDLTSAAMYRGVYALHD-VSDKIPMTNSPILDPLFNLIKIVYNSGSAVTPQDDLAIFS 498  
QY 480 SRLS---TONYF-----RSLPRGT--SNMTYGTENFLGRLMIPNTGISLLIPDDA 525  
DB 499 SKLSPQMTQSLLENEALNKNQSLARQTPDPSCTAFGTNSLGGHLLIPNSGVSLIPAGA 558  
QY 526 IPRGKIYEIYLTLLHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVLAMDHGCEPSPD 585  
DB 559 IPQGRVYEMVTVHRKENMRPMDSTQLLTVPVSCGPPGALLTRPVILTHHCADPSTE 618  
QY 586 SWSRLKKQSCGSEWODVLHGBEAPSHLYYCOLEASACVVFTEQLGRFALVGEALVA 645  
DB 619 DMKIQLKNOAVQGW-E-DVVVVGEEFTTFCYIQLDAEACHILTENLSTVALVQSTTKA 677  
QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKEVVQLEKQLGGQLIQEPRVILHFKDS 705  
DB 678 AAKRLKLAIFGLCCSSLSYIRVYCLDDTDQDALKEVQLEROMGGQLJUEEPKALHFKGS 737  
QY 706 YHNLRLSIHDVPSLSLWKSLLVSYQETPFVHIWNGTORYLHCTFTLERSVSPSTDIAACKL 765  
DB 738 IHNLRSLIHDIAHLSLWKSLLAKYQEI PFVHIWNGSQRLNHLCTFTLERSLNTVELVCKL 797  
QY 766 WYVQVEGDGQSFNFINITKDTFAELLALASEAGVPALVGPFAFKIPFLIROKILISLD 825  
DB 798 CVRQVEGEGQIFQLNCTVSEETGIDPLLDPASTTTTGTGSAFSLPLIRQKLCSSLD 857  
QY 826 PPCRRGADWRTLAKGLHDLSHLGFASPKPSPTAMILNLWEARHPNGLSQAALAAVAGLG 885  
DB 858 APOTRGHWRMLAHKLNLDRLNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMG 917  
QY 886 QPDAGLFTVSEAE 898  
DB 918 RHETTVVSLAAEQ 930

## RESULT 13

US-10-087-684-35  
; Sequence 35, Application US/10087684  
; Publication No. US20040029116A1  
; GENERAL INFORMATION:  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David J.  
; APPLICANT: Grosse, William M.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catharine E.  
; APPLICANT: Casman, Stacie, J.  
; APPLICANT: Spytek, Kimberly A.



[illegible]





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 18:16:33 ; Search time 27 Seconds  
(without alignments)  
1718.954 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGIWPALLGIVLAWL.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4545.5	95.0	898	2	US-08-808-982-5
2	4545.5	95.0	898	3	US-09-306-902A-5
3	2789.5	58.3	557	2	US-08-808-982-6
4	2789.5	58.3	557	3	US-09-306-902A-6
5	2490.5	52.0	943	2	US-08-808-982-7
6	2490.5	52.0	943	3	US-09-306-902A-7
7	294	6.1	102	2	US-08-808-982-8
8	294	6.1	102	3	US-09-306-902A-8
9	273.5	5.7	1172	1	US-08-313-288B-19
10	249	5.2	441	3	US-08-985-526-3
11	246	5.1	239	5	PCT-US93-01652-1
12	246	5.1	1170	1	US-08-313-288B-20
13	235.5	4.9	218	3	US-08-985-526-1
14	234	4.9	1395	3	US-09-540-245A-15
15	216.5	4.5	1651	3	US-09-540-245A-18
16	208.5	4.4	469	1	US-08-313-288B-15
17	205.5	4.3	1381	3	US-09-540-245A-16
18	198.5	4.1	788	2	US-08-918-914-4
19	194.5	4.1	1069	4	US-09-877-730-2
20	194.5	4.1	1150	4	US-09-877-730-8
21	189	3.9	904	4	US-09-877-730-6
22	189	3.9	985	4	US-09-877-730-10
23	188	3.9	1297	3	US-09-540-245A-17
24	186	3.9	380	4	US-09-877-730-4
25	186	3.9	1266	4	US-08-506-296B-4
26	168	3.5	481	4	US-09-130-491-8
27	167.5	3.5	905	4	US-09-369-364A-9

28	166	3.5	837	4	US-09-122-126B-2	Sequence 2, Appli
29	166	3.5	837	4	US-09-634-286A-2	Sequence 2, Appli
30	160.5	3.4	949	4	US-09-568-559-2	Sequence 2, Appli
31	160.5	3.4	967	4	US-09-130-491-2	Sequence 2, Appli
32	159.5	3.3	727	4	US-09-445-023A-1	Sequence 1, Appli
33	159.5	3.3	1224	4	US-09-930-872-4	Sequence 4, Appli
34	158	3.3	608	4	US-09-130-491-13	Sequence 13, Appli
35	156	3.3	997	4	US-09-369-364A-7	Sequence 7, Appli
36	153.5	3.2	1059	4	US-09-907-794A-290	Sequence 290, App
37	153.5	3.2	1059	4	US-09-905-125A-290	Sequence 290, App
38	153.5	3.2	1059	4	US-09-902-775A-290	Sequence 290, App
39	153.5	3.2	1119	4	US-09-907-794A-294	Sequence 294, App
40	153.5	3.2	1119	4	US-09-905-125A-294	Sequence 294, App
41	153.5	3.2	1119	4	US-09-902-775A-294	Sequence 294, App
42	153	3.2	1081	4	US-09-369-364A-17	Sequence 17, Appl
43	152.5	3.2	757	4	US-09-963-791-24	Sequence 24, Appl
44	152.5	3.2	908	4	US-09-963-791-2	Sequence 2, Appli
45	152.5	3.2	1911	1	US-08-348-006B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-08-808-982-5  
; Sequence 5, Application US/08808982  
; Patent No. 5939271  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsey  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,982  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 898 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-08-808-982-5

Query Match 95.0%; Score 4545.5; DB 2; Length 898;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGIWPALLGIVLAWL.....AVAGLGQPDAGLFTVSEAC 899  
|||||

Db 1 MAVRPGMLPVLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPDEVYIVKNKP 60  
QY 61 VLLVCXAVPATQIPFFKNCGEWVRQVDHVIERSIDGSGEPTMEVRINVSQOQEVKVFGL 120  
Db 61 VLLVCXAVPATQIPFFKNCGEWVRQVDHVIERSIDGSGEPTMEVRINVSQOQEVKVFGL 120  
QY 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180  
Db 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180  
QY 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240  
Db 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240  
QY 241 VNGGWSWTWTSVCSACGRGWQKRSCTNPAPNGAFCEGQNVHRTVSSLLVSDG 300  
Db 241 VNGGWSWTWTSVCSACGRGWQKRSCTNPAPNGAFCEGQNVHRTVSSLLVSDG 300  
QY 301 SWSWKSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEV 360  
Db 301 SWSWKSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEV 360  
QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGQVSIKPSKADNPHL 420  
Db 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGQVSIKPSKADNPHL 420  
QY 421 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479  
Db 421 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479  
QY 480 SRLSTQNYFRSLPRGTSNMYGTNFGGRLMIPNTGISLLIPPDALPRGKIYIYTLH 539  
Db 480 SRLSTQNYFRSLPRGTSNMYGTNFGGRLMIPNTGISLLIPPDALPRGKIYIYTLH 539  
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSPDSWLSRLKQSCGS 599  
Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSPDSWLSRLKQSCGS 599  
QY 600 WEQVILHGBAPSHLYYQLEASACVYFTQOLGRFALVGEALSVAARLKLILFAPVA 659  
Db 600 WE-DVLHGEESPSHLYYQLEAGACVYFTQOLGRFALVGEALSVAARLKLILFAPVA 658  
QY 660 CTSLEYNIRVCLDTHDALKEVYQLEKQGLIQEPRVLHFKDSYHNLRLSLHDVPSS 719  
Db 659 CTSLEYNIRVCLDTHDALKEVYQLEKQGLIQEPRVLHFKDSYHNLRLSLHDVPSS 718  
QY 720 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSF 779  
Db 719 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSF 778  
QY 780 NFNTKTQTRFAELLALAESEAGVPALVGSAPKIPFLIROKLISLIDPCREGADWRTIAQ 839  
Db 779 NFNTKTQTRFAELLALAESEAGVPALVGSAPKIPFLIROKLISLIDPCREGADWRTIAQ 838  
QY 840 KLHLDLSLFSFASKPSPTAMILNWEARHFPNGNLSQLAAVAGLQPDAGLFTVSEAE 899  
Db 839 KLHLDLSLFSFASKPSPTAMILNWEARHFPNGNLSQLAAVAGLQPDAGLFTVSEAE 898

## RESULT 2

US-09-306-902A-5

; Sequence 5, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE &amp; TECHNOLOGY LAW GROUP

STREET: 268 BUSH STREET, SUITE 3200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/306.902A

FILING DATE: 07-May-1999

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: UC96-217

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 898 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-306-902A-5

Query Match

Best Local Similarity 95.0%; Score 4545.5; DB 3; Length 898;

Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGMLPVLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPDEVYIVKNKP 60

Db 1 MAVRPGMLPVLGIVLAALRGSGAQOSATVANPVGANPDLLPHFLVEPDEVYIVKNKP 60

QY 61 VLLVCXAVPATQIPFFKNCGEWVRQVDHVIERSIDGSGEPTMEVRINVSQOQEVKVFGL 120

Db 61 VLLVCXAVPATQIPFFKNCGEWVRQVDHVIERSIDGSGEPTMEVRINVSQOQEVKVFGL 120

QY 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180

Db 121 EYWCQCVAMSSSGTTKQKAYIRIARLKNPEQPLAKEVSLGQIVLPCRPPGIPPAE 180

QY 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240

Db 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240

QY 241 VNGGWSWTWTSVCSACGRGWQKRSCTNPAPNGAFCEGQNVHRTVSSLLVSDG 300

Db 241 VNGGWSWTWTSVCSACGRGWQKRSCTNPAPNGAFCEGQNVHRTVSSLLVSDG 300

QY 301 SWSWKSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEV 360

Db 301 SWSWKSACGLDCTHWRSECDPAPRNGEGECQGTDLDRNCTSDLCVHSASGPEV 360

QY 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGQVSIKPSKADNPHL 420

Db 361 ALYVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGQVSIKPSKADNPHL 420

QY 421 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479

Db 421 LTIQPDLS-TTTTYQGSICPRQDQSPKQFQTNHGLLSPLGGGRHTLHSSPTSEAEFV 479

QY 480 SRLSTQNYFRSLPRGTSNMYGTNFGGRLMIPNTGISLLIPPDALPRGKIYIYTLH 539

Db 480 SRLSTQNYFRSLPRGTSNMYGTNFGGRLMIPNTGISLLIPPDALPRGKIYIYTLH 539

QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSPDSWLSRLKQSCGS 599

Db 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSPDSWLSRLKQSCGS 599

540	KP	DVRLPLACQ	ILLSPVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRKKQSC	EGS	599
600	WE	QDVLHLGEE	APSHLYCYQLEASACYVFTQ	LG	659
600	WE	DVLHLGEE	SPSHLYCYQLEAGACYVFTQ	LG	658
660	CTS	LEYNIRVY	CLDHTDHALKEVVQLEKQ	LG	719
659	CTS	LEYNIRVY	CLDHTDHALKEVVQLEKQ	LG	718
720	LW	SKLLVSQY	QEIIFYHIWNGTQRYLHCTFT	LER	779
719	LW	SKLLVSQY	QEIIFYHIWNGTQRYLHCTFT	LER	778
780	NF	NTKOTR	FAELALRESEAGVPALVGFSA	KI	839
779	NF	NTKOTR	FAELALRESEAGVPALVGFSA	KI	838
840	KL	HLDSHLS	FPASKPSTAMILNLWEARH	FP	899
839	KL	HLDSHLS	FPASKPSTAMILNLWEARH	FP	898

RESULT 3  
 US-08-808-982-6  
 ; Sequence 6, Application US/08808982  
 ; Patent No. 5939271  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tessier-Lavigne, Marc  
 ; APPLICANT: Leonardo, E. David  
 ; APPLICANT: Hink, Lindsay  
 ; APPLICANT: Masu, Masayuki  
 ; APPLICANT: Kazuko, Keino-Masu  
 ; TITLE OF INVENTION: Netrin Receptors  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 268 BUSH STREET, SUITE 3200  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/808,982  
 ; FILING DATE:  
 ; CLASSIFICATION: 530  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OSMAN, RICHARD A.  
 ; REGISTRATION NUMBER: 36,627  
 ; REFERENCE/DOCKET NUMBER: UC96-217  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 343-4341  
 ; TELEFAX: (415) 343-4342  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 557 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: peptide  
 ; US-08-808-982-6

Query Match	58.3%	Score	2789.5	DB 2	Length	557			
Best Local Similarity	96.4%	Pred. No.	2,7e-256						
Matches	538	Conservative	2	Mismatches	15	Indels	3	Gaps	3

344 NCTSDLCVHSASPEDVALYVGLIAVAVCLVLLVLLVILVYCRKEGLSDSDVADSSILTS 403

Db	1	NCTSDLVHTASGDPEDVALYVGLLIAVAVCLVLLLVLLVILVYCRKKEGLDSDVADSSILTS	60
QY	404	GFOVSVTKPSKADNPHLITTOPDLIS- TTTTYYQSLCPRDQSPKXQLTNGHLLSPGLGG	462
Db	61	GFOVSVTKPSKADNPHLITTOPDLIS- TTTTYYQSLCPRDQSPKXQLTNGHLLSPGLGG	120
QY	463	RHTLHSSPTSEAEFEFVSRSLSTQNYFRSLPRGTSNMVYGTNFNLGRMLMNPWTGSLILP	522
Db	121	RHTLHSSPTSEAEFEFVSRSLSTQNYFRSLPRGTSNMVYGTNFNLGRMLMNPWTGSLILP	180
QY	523	PDAPRGKIYEIYLTLLHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP	582
Db	181	PDAPRGKIYEIYLTLLHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP	240
QY	583	SPDSWSLRLKKQSCGSEWODVLHGREAPSHLYCOLEASACYVFTBQLGRFALVGEAL	642
Db	241	SPDSWSLRLKKQSCGSEWODVLHGREAPSHLYCOLEASACYVFTBQLGRFALVGEAL	299
QY	643	SVAAAKELKLLLPAPVACTSLEYNIRVYVCLDTHDALKVVOLEKQGLQIQEPRVLHF	702
Db	300	SVAAAKELKLLLPAPVACTSLEYNIRVYVCLDTHDALKVVOLEKQGLQIQEPRVLHF	359
QY	703	KDSYHNRLRSLHDVPSLSKLLVSYOEIPFYHIWNGTORYLHCTFTLERSVSPSTDIA	762
Db	360	KDSYHNRLRSLHDVPSLSKLLVSYOEIPFYHIWNGTORYLHCTFTLERSVSPSTDIA	419
QY	763	CKLWVQVEGDGQSFSEFNITKDTFRAELIALESEAGVPALVGPFAKIPFLIROKIIS	822
Db	420	CKLWVQVEGDGQSFSEFNITKDTFRAELIALESEAGVPALVGPFAKIPFLIROKIIS	479
QY	823	SLDPPCRRGADWTLAQKHLDSHLSFFASKPSPMTLNLWEARHPNGNLSOLAAVA	882
Db	480	SLDPPCRRGADWTLAQKHLDSHLSFFASKPSPMTLNLWEARHPNGNLSOLAAVA	539
QY	883	GLQOPDAGLFT-VSEAC 899	
Db	540	GTXPAGRWLLSQCSEAC 557	
RESULT 4			
US-09-306-902A-6			
; Sequence 6, Application US/09306902A			
; Patent No. 6277585			
; GENERAL INFORMATION:			
; APPLICANT: Tessier-Lavigne, Marc			
; Leonardo, E. David			
; Hink, Lindsay			
; Masu, Masayuki			
; Kazuko, Keino-Masu			
; TITLE OF INVENTION: Netrin Receptors			
; NUMBER OF SEQUENCES: 9			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP			
; STREET: 268 BUSH STREET, SUITE 3200			
; CITY: SAN FRANCISCO			
; STATE: CALIFORNIA			
; COUNTRY: USA			
; ZIP: 94104			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/306,902A			
; FILING DATE: 07-May-1999			
; CLASSIFICATION: <Unknown>			
; ATTORNEY/AGENT INFORMATION:			
; NAME: OSMAN, RICHARD A			
; REGISTRATION NUMBER: 36,627			
; REFERENCE/DOCKET NUMBER: UC96-217			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (415) 343-4341			

```

Query Match          58.3%; Score 2789.5; DB 2; Length 557;
Best Local Similarity 96.4%; Pred. No. 2,7e-256;
Matches 538; Conservative 2; Mismatches 15; Indels 3; Gaps 3;

QY      344 NCTSDLCVHSASPEDVALYVGLIAVAVCLVLVLLVLLVLYCRKKEGLSDSDVADSSILTS 403
      |||:|||||

```

TELEFAX: (415) 343-4342  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 557 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 US-09-306-902A-6

Query Match 58.3%; Score 2789.5; DB 3; Length 557;  
 Best Local Similarity 96.4%; Pred. No. 2.7e-256;  
 Matches 538; Conservative 2; Mismatches 15; Indels 3; Gaps 3;  
 344 NCTSDLCVHSGASPEDVALVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTS 403  
 1 NCTSDLVHTASGEDVALVGLIAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTS 60  
 404 GFQPVSIKPKADNPHTLTTPDLS-TTITYQGLCPQDQSPKFKLTNGHLLSPGGG 462  
 61 GFQPVSIKPKADNPHTLTTPDLS-TTITYQGLCPQDQSPKFKLTNGHLLSPGGG 120  
 463 RHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGFNFGGLRMLPNTGISLLIP 522  
 121 RHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGFNFGGLRMLPNTGISLLIP 180  
 523 PDAIPRGKIYEIYLTTHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP 582  
 181 PDAIPRGKIYEIYLTTHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP 240  
 583 SPDSWSRLKQSCGEGWEQDVLHGBEAPSHLYYQLEASACVFTQLGRFALVGEAL 642  
 241 SPDSWSRLKQSCGEGWE-DVLHGBEAPSHLYYQLEASACVFTQLGRFALVGEAL 299  
 643 SVAARKLKLFPAPVACTSLEYNIRVYCLHDTDALKEVVOLEKQGLIQEPRVLHF 702  
 300 SVAARKLKLFPAPVACTSLEYNIRVYCLHDTDALKEVVOLEKQGLIQEPRVLHF 359  
 703 KDSYHNRLSLTHDVPSSLSKLVSYQIEPFYHWNQYRQLHCTFLERSVPSSTDLA 762  
 360 KDSYHNRLSLTHDVPSSLSKLVSYQIEPFYHWNQYRQLHCTFLERSVPSSTDLA 419  
 763 KLVWVQVEGQGSFNFNTKTQTRAEALLALESEAGVPALVGSAPFKIFLIRQKLIIS 822  
 420 KLVWVQVEGQGSFNFNTKTQTRAEALLALESEAGVPALVGSAPFKIFLIRQKLIIS 479  
 823 SLDPCCRGADWRTLAOKLHLDHSLSPFASKPSPTAMILNLWEARHPFNGLSQLAAVA 882  
 480 SLDPCCRGADWRTLAOKLHLDHSLSPFASKPSPTAMILNLWEARHPFNGLSQLAAVA 539  
 883 GLGQPDAGLFT-VSEAE 899  
 540 GTXPAGRWLLSQCEAE 557

RESULT 5  
 US-08-808-982-7  
 Sequence 7, Application US/08808982  
 Patent No. 5939271  
 GENERAL INFORMATION:  
 APPLICANT: Tessier-Lavigne, Marc  
 APPLICANT: Leonardo, E. David  
 APPLICANT: Hink, Lindsey  
 APPLICANT: Masu, Masayuki  
 APPLICANT: Kazuko, Keino-Masu  
 TITLE OF INVENTION: Netrin Receptors  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 BUSH STREET, SUITE 3200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA

COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA: US/08/808,982  
 APPLICATION NUMBER: US/08/808,982  
 FILING DATE:  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OSMAN, RICHARD A.  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: UC96-217  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 343-4341  
 TELEFAX: (415) 343-4342  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 943 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: peptide  
 US-08-808-982-7  
 Query Match 52.0%; Score 2490; DB 2; Length 943;  
 Best Local Similarity 52.3%; Pred. No. 2e-227;  
 Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;  
 9 PALLGIVLAWLKSGGAQSSATVANPVGANPDLLPHLEVPEDVYLVKNKPVLLVCKAV 68  
 21 PSLAGI-----DSGAQ---GLPDSFSPAPAEQLPHLEPEDAYLVKNKPVLEHCRAF 70  
 69 PATQIFKFCNGEWRQVDHVIERTDGSSEPTMEVRINVSROQVKEVFGLEEYWCQVA 128  
 71 PATQIFKFCNGEWSQKHVTVQESLDEATGLRREVQIEVSROQVVEELFGLEDYWCQVA 130  
 129 WSSGTTKQKAYIRIARLNKFNQEPLEKVEQLGIVLPCRPEGIIPPAEVEWLRNED 188  
 131 WSSGTTKSRAYIRIARLNKFNQEPLEKVEQLGIVLPCRPEGIIPPAEVEWLRNED 190  
 189 LVDSLPDNPVYITEHSLVVRQARLADTANTVCAKNIVARRRSASAIVVYVNGGSTW 248  
 191 VIDPAQDTNFLTIDHNLIIQARLSDTANTVCAKNIVARRRSASAIVVYVNGGSTW 250  
 249 TEWSVCSASCGRWQKRSRCTNPAPLNGGAFCEGQNVHRTVSSLVSDVGSWSPWSKW 308  
 251 AEWSPCSNRCGRGWQKRTCTNPAPLNGGAFCEGQ-ACQKTACTTVCVDGAWTEWSKW 309  
 309 SACGLDCTHWRSECSDPAPRNGEBCQGTDLDRNCTSDLCV---HSASGE----- 358  
 310 SACSTECARHWRSECMAPPQNGGRDCSTGLLDSKNCTDGLCVLNQRTLNDRPKSRPLEPS 369  
 359 -DVALYVGL-IYAVAVCLVLLVLLVLYCRKKEGLSDVADSS-ILTSGFQPVSIKPSKA 415  
 370 GDVALYVGLVAVFVVLAVLMAVGVIYRRNCRDFDITDSSAALGGFHPVFKTARP 429  
 416 DNPILL-TIQPDLSTTT-TYQGLCPQDQSPKFKLTNGHLLSP----- 459  
 430 SNPQLHPSAPDLTASAGIYRGPVYALQDS-ADKIPMTNSPFLDPLSLKIVYDSSTI 488  
 460 ---GGG-----RHTLHSSPTSEAEFVSRLSTQNYFRSLPRG 494  
 489 GSGAGLADGADLLGLVPPGTYPGDFSRTHFLHRS-----ASLGSQ-HLLGLPRD 538  
 495 TSNMTYGFNFGGLRMLPNTGISLLIPDAPRGKIYEIYLTTHKPEDVRLPLA-GCQT 553  
 539 PSSSVSGTFCGLGRLTIPGTGVSLLVNGAIPOGKFYDLYLRINKTEST-LPLSEGSQT 597  
 554 LLSITVSCGPPGVLLTRPVILAMDHCGEPSPDSWSRLKQSCGEGWEQDVLHGBEAP 613

Db 598 VLSFVTCGPTGLLLCRPVLVTPHCAEVIAGDWIFOLKTAHQHWE-EVVTLDDETLN 656  
Qy 614 HLYYCOLEASACVYFTEOLGRFALVGAALVAARLKLKLLFAPVACTSLEYNIRVYCLH 673  
Db 657 TPCYCOLEAKSCHILLDOLGYVFTGESYSRAVKRLQLAIFAPALCTSLSEYLRVYCLE 716  
Qy 674 DTHDALKEVVQLEKQGLIQEPVHLHFQDSYHNLRSLSDHVPSSLWKSLLVSYOEIP 733  
Db 717 DTPAALKEVLELERTLGGYLVEEPKTLFKDSYHNLRSLSDHDPHAWRSKLLAKYOEIP 776  
Qy 734 FYHWNQTOYRLHCTFTLERVSPSTSLACKLWVQVGEQGSFINSFNITKOTRFAELL 793  
Db 777 FYHWNQSKALHCTFTLERHSLASTFTCKVCVRQVEGEQIFQLHTTILA-ETPAGSLD 835  
Qy 794 ALESEAGVPA--VGSAPKIPFLIPIROKICNSLDAPNSRGNDWRLLAQKLSMDRYLYNFA 851  
Db 836 ALCSAPGNAATTOLGYPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLYNFA 895  
Qy 852 SKPSPTAMILNWEARHPFNGLSQAALAAVAGLQPDAGLFTVSEAE 899  
Db 896 TKASPTGVILDLWEARQDDGLNSLASALEEMCKSEMLVAMTTDGD 943

## RESULT 6

US-09-306-902A-7

; Sequence 7, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsey

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE &amp; TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/306,902A

; FILING DATE: 07-May-1999

; CLASSIFICATION: &lt;Unknown&gt;

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC96-217

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 943 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-306-902A-7

Query Match

Best Local Similarity 52.0%; Score 2490; DB 3; Length 943;

Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;

Qy 9 PALLGIVLAALWLRGGAQQAQATVANPVGANPDLLPHFLVEPVDVYVKNKPVLLVCKAV 68

Db 21 PSLAGI-----DSGAQ---GLPDSFSPAPAEQLPHLEPEDAYIVKNKPVLLHCRAF 70  
Qy 69 PATQIFPKONGEWYRQVDHVIERTDSSGSEPTMEVRINVSQVQKVFGLGEYWCQVCA 128  
Db 71 PATQIFKNGEWYSQKGVHTQESLDEATGLRIREVQIEVSQVQVEELFGLGEYWCQVCA 130  
Qy 129 WSSSGITKSQKAYIRIARLKNFROBLAKEVSLVEQGIIVLPCRPEGIPPAEVEWLNRN 188  
Db 131 WSSSGITKSRRAYIRIAYLRKNFQDEPLAKEVPLDHEVLLQCRPEGPGVPAEVEWLKNE 190  
Qy 189 LVDSLDPNVYITREHSLVVRQARLADANTVCAKNIVARRRSASAAVYVNGSWSTW 248  
Db 191 VIDPAQDTNELLTDHNLIIROARLSDTANTVCAKNIVARRRSATATVYVNGSWSSW 250  
Qy 249 TEMSVCSASCGRWKQKRSRSTNPAPLNGGAFCEGQNVHRTVSSLLVSDGSMSPKSW 308  
Db 251 AEWSPCSNRGCGWKQKTRCTNPAPLNGGAFCEGQ--ACQKTACTTVCVVDGAWTEWSK 309  
Qy 309 SACGLDCTHWSRECSDPAPNGGECQGTDLTRNCTSDLCV---HSAGSPE----- 358  
Db 310 SACSTECARHSRECMAPPQNGRCDCGTLSDSKNCTDGLCVLNQRTLNDPKSRPLEPS 369  
Qy 359 -DVALYVGL-IYAVAVCLVLLLVLYCRKKEGLSDSDVADSS-ILTSGPQVSIKESKA 415  
Db 370 GDVALYAGLVAVFVVLAVLMVAGVIVRRNCRDFTDITDSSAALTGGFHPVNFKTARP 429  
Qy 416 DNPHLL--TIQPDLSITT--TYQGSCLPRQDGPSPKFQLTNGHLLSPL----- 459  
Db 430 SNPQLLHPSAPPDLTASAGIYRGPVVALQDS-ADKIPTMTNSPLLDPLSLKIVYDSTI 488  
Qy 460 --GGG-----RUTLHSSPTSEAEFVSRSLSTQNYPRSRPRG 494  
Db 489 GSGAGLADGADLLGLVPPGTYPGDFSRDTHFLRS-----ASLSGQ-HLLGLPRD 538  
Qy 495 TSNNYTGTFNFGRLMIPTGISLILPPDAIPRGKIYEIVLTHKPEDVRLPLA-CCOT 553  
Db 539 PSSSVSTFCGLGRITPCTGVSLVVPNGAIFQKGYDLYLRINKTEST-LPISEGSQT 597  
Qy 554 LLSPIVSCGPPGVLLTRPVILAMDHCCEPSPDSMSLRLLKQSCGSGWEQDVLHGEAPS 613  
Db 598 VLSFVTCGPTGLLLCRPVLVTPHCAEVIAGDWIFOLKTAHQHWE-EVVTLDDETLN 656  
Qy 614 HLYYCOLEASACVYFTEOLGRFALVGAALVAARLKLKLLFAPVACTSLEYNIRVYCLH 673  
Db 657 TPCYCOLEAKSCHILLDOLGYVFTGESYSRAVKRLQLAIFAPALCTSLSEYLRVYCLE 716  
Qy 674 DTHDALKEVVQLEKQGLIQEPVHLHFQDSYHNLRSLSDHVPSSLWKSLLVSYOEIP 733  
Db 717 DTPAALKEVLELERTLGGYLVEEPKTLFKDSYHNLRSLSDHDPHAWRSKLLAKYOEIP 776  
Qy 734 FYHWNQTOYRLHCTFTLERVSPSTSLACKLWVQVGEQGSFINSFNITKOTRFAELL 793  
Db 777 FYHWNQSKALHCTFTLERHSLASTFTCKVCVRQVEGEQIFQLHTTILA-ETPAGSLD 835  
Qy 794 ALESEAGVPA--VGSAPKIPFLIPIROKICNSLDAPNSRGNDWRLLAQKLSMDRYLYNFA 851  
Db 836 ALCSAPGNAATTOLGYPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAQKLSMDRYLYNFA 895  
Qy 852 SKPSPTAMILNWEARHPFNGLSQAALAAVAGLQPDAGLFTVSEAE 899  
Db 896 TKASPTGVILDLWEARQDDGLNSLASALEEMCKSEMLVAMTTDGD 943

## RESULT 7

US-08-808-982-8

; Sequence 8, Application US/08808982

; Patent No. 5932271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsey

; APPLICANT: Masu, Masayuki

```

; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; QUERY MATCH
; Best Local Similarity 6.1%; Score 294; DB 2; Length 102;
; Matches 57; Conservative 16; Mismatches 28; Indels 0; Gaps 0;
;
; Qy 609 EEAPSHLYYQLEASACVFTQIGRPAALVGEALSVAAAKRLKLLLPAPVACTSLEYNIR 668
; Db 2 EETLNTPCYXQLEPRACKILLDQGTGYFTGESYSRAVKRLQLAVFAPALCTSLYSLSR 61
;
; Qy 669 VYCLHDTHDALKEVVQLEKQGGOLIOEPRVLHFKDSYHNL 709
; Db 62 VYCLEDPVALKEVLELRTIGGYLVEEPKPLMPKDSYHNL 102
;
; RESULT 8
; US-09-306-902A-8
; Sequence 8, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindeay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
; US-09-306-902A-8
;
; Query Match 6.1%; Score 294; DB 3; Length 102;
; Best Local Similarity 56.4%; Pred. No. 4.1e-20;
; Matches 57; Conservative 16; Mismatches 28; Indels 0; Gaps 0;
;
; Qy 609 EEAPSHLYYQLEASACVFTQIGRPAALVGEALSVAAAKRLKLLLPAPVACTSLEYNIR 668
; Db 2 EETLNTPCYXQLEPRACKILLDQGTGYFTGESYSRAVKRLQLAVFAPALCTSLYSLSR 61
;
; Qy 669 VYCLHDTHDALKEVVQLEKQGGOLIOEPRVLHFKDSYHNL 709
; Db 62 VYCLEDPVALKEVLELRTIGGYLVEEPKPLMPKDSYHNL 102
;
; RESULT 9
; US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

US-08-313-288B-19

Query Match 5.7%; Score 273.5; DB 1; Length 1172;  
 Best Local Similarity 28.9%; Pred. No. 2.1e-16;  
 Matches 74; Conservative 27; Mismatches 110; Indels 45; Gaps 8;  
 QY 209 RQARLADTANYTCAKNTVARRRSASAA-VIVYVNGWSTWTSVCSASCGRGWQKESR 267  
 DB 403 QRGSRCDVTSNCLGFSIQACSLSKCDTRIQDGGWHSWSPSSCSVTGCVGNITIR 462  
 QY 268 SCTNPAPLNGAGAFCEGQNVHDTVSSLLVSDGWSMWSKWSACGLDCT---HWSRSCS 324  
 DB 463 LCNSPVPQMGKGNCKGSKGRETAKQAGAPCDIGRWSWSPSACTVTCAGGIRERTVCN 522  
 QY 325 DPARNGEEECQGTDLDRNCTSDLCVHSASGPDVALYVGLIAVAVCLVLLVLLVY 384  
 DB 523 SPEFYGGKACGVQERQMCNKRSC-----PVDGCLSNPCFFGAQC----- 564  
 QY 385 CRKKEGLSDVADSSILTSFGQPVSI--KPSKADNPHELLTIQDPLSTTT-----TYQ 434  
 DB 565 -----SSFPDGS-WSCGFCVGFGLNGTHCEDLDECALVPDICTSKVPCVNTQP 615  
 QY 435 GSLC-----PRODGPSP 446  
 DB 616 GFHCLPCPPRYRGNOQ 631

RESULT 10

US-08-985-526-3  
 ; Sequence 3, Application US/08985526  
 ; Patent No. 6080728  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mixson, James A  
 ; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA  
 ; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE  
 ; TITLE OF INVENTION: THERAPY  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Connolly, Bove, Lodge, & Hutz  
 ; STREET: 1220 Market Street, P.O. Box 2207  
 ; CITY: Wilmington  
 ; STATE: Delaware  
 ; COUNTRY: U.S.A.  
 ; ZIP: 19899  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/985,526  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/608,845  
 ; FILING DATE: 16-JUL-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McMorrow Jr., Robert G  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (302) 658-9141  
 ; TELEFAX: (302) 658-5613  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 441 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-985-526-3

Query Match 5.2%; Score 249; DB 3; Length 441;  
 Best Local Similarity 25.9%; Pred. No. 9e-15;  
 Matches 87; Conservative 35; Mismatches 114; Indels 100; Gaps 15;  
 QY 75 FKCNQEW-----VRQVDHVIERTDGSSEPTM-----EVRINVSQQ---V 113

DB 132 FRQDGGWHSWSPSSCSVTCDGVTITLTCNSPQMGKPCGEARETKACKKDACPI 191  
 QY 114 EKVFGLBEYWCQCWAMSSSGTTKQKAXIRIARLRKNPEQPLAKEVSLQGIIVLPCRPP 173  
 DB 192 NGWGPSPWDICSVTCGGGVQKESRLCV---DSRMTTEENKELANELR-----RPP 239  
 QY 174 ----EGIPPAEVEWLNRD-LVDPSLDPNVVITREHSLVVRQARLADTANYTCAKNTVA 228  
 DB 240 LCYHNG-----VOYRNNEEWTSDCTE-----CHCQNSVT 269  
 QY 229 RRSASAAVIVYVNG-----GWSWTWTSVCSASCGRGWQKRSRSC 269  
 DB 270 ICKKVCPIPCSNATVPDGECCPRCWPSDSADDDGWSFWSWTSCSTSCNGIQOGRSC 329  
 QY 270 TNPAFLNGGAFCEQNVHDTV-----SSLLVSDGWSMWSKWSACGLDC---THWSRE 322  
 DB 330 DS---LNNR--CEGSSVQTRTCHIQCCKRQKQGGHSHWSPSSCSVTCDGVTITL 384  
 QY 323 CSDPAPRNGEEECQGTDLDRNCTSDLC-VHSASGP 357  
 DB 385 CNSPQMGKPCGEARETKACKKDACPINGWGP 420

RESULT 11

PCT-US93-01652-1  
 ; Sequence 1, Application PC/TUS9301652  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bouck, Noel P.  
 ; APPLICANT: Polverini, Peter J.  
 ; APPLICANT: Good, Deborah J.  
 ; APPLICANT: Frazier, William A.  
 ; TITLE OF INVENTION: Method and Composition for  
 ; TITLE OF INVENTION: Inhibiting Angiogenesis  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Filton, Fallon, Lungmus & Chestnut  
 ; STREET: 100 South Wacker Drive, Suite 960  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606-4002  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US93/01652  
 ; FILING DATE: 19930222  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/841,656  
 ; FILING DATE: 24-FEB-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/464,369  
 ; FILING DATE: 12-JAN-1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fentress, Susan B.  
 ; REGISTRATION NUMBER: 31,327  
 ; REFERENCE/DOCKET NUMBER: 92005-PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (312)-456-8000  
 ; TELEFAX: (312)-456-7776  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 239 amino acids  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS: unknown  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 ; PCT-US93-01652-1

Query Match 5.1%; Score 246; DB 5; Length 239;  
 Best Local Similarity 31.0%; Pred. No. 6.2e-15;  
 Matches 48; Conservative 23; Mismatches 66; Indels 18; Gaps 3;

QY 207 VVQRLADTANYTCVAKNIVAR-----RRSASAAVIVYNGGWSTWTSVCSASC 258  
 DB 88 IQQGRSCDSLNRCGSSVQTRTCHIQCDKRFKQ-----DGGSHWSPWSSCSVTC 140

QY 259 GRGQKRSCTNPAPLNGAFCEGQNVHRTVSSLLVSDGSPWSPWKSACGLDC--- 315  
 DB 141 GDGVITRLCNLSPQWGLPCEGEARETKACKKDACPINGGWGSPWMDICSVTCGGG 200

QY 316 THWSRECSDPAPRNGGECQGTDLTDRNCTSDLC 350  
 DB 201 VQKSRLCNNPAPQFGGLDCVGVDTENQICNKQDC 235

RESULT 12  
 US-08-313-288B-20  
 ; Sequence 20, Application US/08313288B  
 ; Patent No. 5750502  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jessell, Thomas M. and Avihu Klar  
 ; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
 ; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/313.288B  
 ; FILING DATE: January 5, 1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-313-288B-20

Query Match 5.1%; Score 246; DB 1; Length 1170;  
 Best Local Similarity 30.3%; Pred. No. 8.8e-14;  
 Matches 47; Conservative 24; Mismatches 66; Indels 18; Gaps 3;

QY 207 VVQRLADTANYTCVAKNIVAR-----RRSASAAVIVYNGGWSTWTSVCSASC 258  
 DB 399 IQQGRSCDSLNRCGSSVQTRTCHIQCDKRFKQ-----DGGSHWSPWSSCSVTC 451

QY 259 GRGQKRSCTNPAPLNGAFCEGQNVHRTVSSLLVSDGSPWSPWKSACGLDC--- 315  
 DB 452 GDGVITRLCNLSPQWGLPCEGEARETKACKKDACPINGGWGSPWMDICSVTCGGG 511

QY 316 THWSRECSDPAPRNGGECQGTDLTDRNCTSDLC 350

Db 512 VQKSRLCNNPTQFGGKDCVGVDTENQICNKQDC 546

RESULT 13  
 US-08-985-526-1  
 ; Sequence 1, Application US/08985526  
 ; Patent No. 6080728  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mixson, James A  
 ; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA  
 ; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE  
 ; TITLE OF INVENTION: THERAPY  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Connolly, Bove, Lodge, & Hutzel  
 ; STREET: 1220 Market Street, P.O. Box 2207  
 ; CITY: Wilmington  
 ; STATE: Delaware  
 ; COUNTRY: U.S.A.  
 ; ZIP: 19899  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/985.526  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/608,845  
 ; FILING DATE: 16-JUL-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McMorriow Jr., Robert G  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (302) 658-9141  
 ; TELEFAX: (302) 658-5613  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 218 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-985-526-1

Query Match 4.9%; Score 235.5; DB 3; Length 218;  
 Best Local Similarity 38.5%; Pred. No. 5.3e-14;  
 Matches 47; Conservative 16; Mismatches 46; Indels 13; Gaps 5;

QY 244 GWSTWTSVCSASCGRGWQKRSCTNPAPLNGAFCEGQNVHRTV---SSLLVSVD 299  
 DB 81 GWSWSEWTSCTSCGNGIQGRSCDS---LNNR--CEGSSVQTRTCHIQCDKRFKQD 135

QY 300 GWSWSPWKSACGLDC---THWSRECSDPAPRNGGECQGTDLTDRNCTSDLC-VHSAS 355  
 DB 136 GWSWSPWSSCSVTCGVDVITRLCNLSPQWGLPCEGEARETKACKKDACPINGGW 195

QY 356 GP 357  
 DB 196 GP 197

RESULT 14  
 US-09-540-245A-15  
 ; Sequence 15, Application US/09540245A  
 ; Patent No. 6270984  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Goodman, Corey  
 ; APPLICANT: Kid, Thomas  
 ; APPLICANT: Brose, Katja  
 ; APPLICANT: Tessier-Lavigne, Marc  
 ; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
 ; FILE REFERENCE: B98-031-3  
 ; CURRENT APPLICATION NUMBER: US/09/540.245A



; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 60/065,544  
; PRIOR FILING DATE: 1997-11-14  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 1395  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-540-245A-15

Query Match 4.9%; Score 234; DB 3; Length 1395;  
Best Local Similarity 20.5%; Pred. No. 1.6e-12;  
Matches 188; Conservative 103; Mismatches 262; Indels 364; Gaps 46;

Qy 4 RPGLWPALLGVLAAWLGSGAQSATVANVPGA-----NPDLPLHFLVPEDEVYVKN 58  
Db 28 RMWLLPAWLLVLA-----SNGLPVRGQYQSPRIEH-----PTDLVVKKN 70  
Qy 59 KPVLLVCK--AVPATQIFKCKNGEVRQVDHVIERSIDGSGEPTMEVRINVR----- 110  
Db 71 EPATLNCKVEKPEPTI-----EWFK-----DGEF-----VSTNEKKSHRVQF 108  
Qy 111 -----QQVEKVFLEBYWCQVAMSSSGTTKQKAYIRIARLKNFQEQPLAKEV 160  
Db 109 KDGALFFVRTWQKKEQDGGYW--CVAKNRVGOAVSRHASLQIAVLDRDFRVEPKTRV 166  
Qy 161 SLEGIVLPCBPPIGPIPAEVEWURN-----EDLVDSL--DPNYYITREHSLVVRQARLA 214  
Db 167 AKGETALLECGPKPIGPEPTLIMIKDGVPLDLDKAMSGFASRRVRIVDGNNLLISNVEPI 226  
Qy 215 DTANYTCVAKNIVARRRSASAAVLYVN-----GWSITWTEWSVCSACGR 260  
Db 227 DEGNKICIAQNLVGTRESSYAKGLIVQKPYFMKEPKDQVMLYGQTATF-----HCSV 278  
Qy 261 G-----WQK-----RSRSCINPAPLNGGAF--CEGQN----- 285  
Db 279 GGDPPPKVLWKKEBGNIPVSRARILHDEKLSLEISNITPTDEGTVCCHANNVGOISARAS 338  
Qy 286 --VH-----DRTVSSLLVSDG-----SWSP-----WSKWSACGL--DCTHWSRE 322  
Db 339 LIVHAPFNFKRPSNKKVGLGVVQLPCMASGNPPSPVFWTKEGVTLMPFNSSHGRQYV 398  
Qy 323 CSDPAPRNGEECOGTLDRNCTSDLCVHSASGPEDEVYVGLIYAVACVILVLLVLL 382  
Db 399 AADGT-----LQITDV-----RQDEGGY-----VCSAF----- 422  
Qy 383 VYCRKKEGLSDVADSSILTSGFPVSIKPSKADN--PHLLTIQPLDSTTTTQGSLL-- 437  
Db 423 -----SVVDSSTVR-----VFLQVSSVDERPPIIIQGP--ANQTLPGSVATL 464  
Qy 438 -CPRQDGPSPKFQ-LTNGHLLSLPGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGT 495  
Db 465 PCRAATGNPSRIKFWFHDHQAQV--AGNRYGIIQGSLLRVDDLQSLDSGTCTASGERGE 522  
Qy 496 SNMTYGTGFLGRLMIPNTGISILLIPDAIPRGIYIYVTLHKPEDEVPLPLAGCOTLL 555  
Db 523 TS-----WAATLVEKPGSTSLHRAA----- 543  
Qy 556 SPITVSGPPGVLITRPIVILAMDHCBPSPDWSLR-LKQSCSGS----- 599  
Db 544 DPSTYPAPGT-----PKVLNV-----SRISILRWAKSKEKPGAVGPIIYTVVEFSPD 593  
Qy 600 -----WEQDVLHGEAPSHIYYCOLEASACYVF-----TEQLGR 634  
Db 594 LQGWIVAAHRVGD---TQVTISGLTPGTSYFVLVRAENTQGISVPSGLSNVIKTIADP 650  
Qy 635 FALVGEALSVA---AAKRLKLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLG 690  
Db 651 DAASANDLSAARTLLTGKSVELIDASAINASAVRLE---WMLHVSAD-----BKYVE 699

Qy 691 GQLICEPRVLHFKDS-----YHNLRL-----SIHDV 716  
Db 700 GLRI-----HYKASVPSAQYHSITVMDASAEFVWGNLKKYTKYEFELTPFFETIEGQ 753  
Qy 717 PSSLWKSXLIVSYOEIP 733  
Db 754 PSN---SKTALTYEDVP 767

RESULT 15  
US-09-540-245A-18  
; Sequence 18, Application US/09540245A  
; Patent No. 6270984  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brosse, Katja  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/540,245A  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 60/065,544  
; PRIOR FILING DATE: 1997-11-14  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 1651  
; TYPE: PRT  
; ORGANISM: human  
US-09-540-245A-18

Query Match 4.5%; Score 216.5; DB 3; Length 1651;  
Best Local Similarity 20.7%; Pred. No. 1e-10;  
Matches 213; Conservative 139; Mismatches 381; Indels 295; Gaps 54;

Qy 41 DLPLHFLVPEDEVYVKNKPVLLVCKAV--PATQIFPKCNGEWWVRQVDHVIERSIDGSSG 98  
Db 65 DFPPRIVEHSDPLSVSGEPATLNCKAEGRPPTI-----EWYKGERVETDKDDPRSH 118  
Qy 99 EPTME-----VRINVSROQVEKVFLEBYWCQVAMSSSGTTKQKAYIRIARLKNFE 152  
Db 119 RMLPSGSLFFLLIVHGRKSRP-----DEGVYCVARNYLGEAVSHNASLEVALRDDR 173  
Qy 153 QEPLAKEVSLQGIPLPCRPPEGIPPAEVEWELRNEDLVDPSPNVYIT--REHSLVVRQA 211  
Db 174 QNPSDVNVAVGEPAVMECQPPRGHPPTISWKKDGSPLD--DKDERITIRGGKLMITYT 230  
Qy 212 RLADTANYTCVAKNIVARRRSASAAVLYVNGW-----STWTEWSV--CSASCGRG 261  
Db 231 RKSDAGKYVCVTGNTMVGERESEVAELTVLERPSFVKRPSNLAVTVDDSAEFKCEA---RG 287  
Qy 262 -----WQK-----RSRSCINPAPLNGGAFCEGQNVHDRTVSSLLV 296  
Db 288 DPVPTVRNRKDGELPKSRYEIRDDHTLKRKYT--AGDMGSYTCVAENMVGKAEASATL 345  
Qy 297 SVDGS-----WSPMSKWSACGLDCTHWRSECSDPAP-----RNGGEEC-----OQTDL 340  
Db 346 TVQEPHFVVKPEPDQVVALGRVT--FQCEATGNPQPAIFWRREGSQNLLFSYOPPOSSSR 404  
Qy 341 DTRNCTSDLCVHSASGPEDEVYVGLIYAVACVILVLLVLLVYCRKKEGLSDVADSSI 400  
Db 405 FSVSQTGDLTIITNVQ--RSDVGYV-----ICQTL-----NVA-GSI 437  
Qy 401 LSGGFPQVSIKPSKADNPHLLTIQPLDSTTTTQGSLL--CPRQDGPSPKF----- 448  
Db 438 ITKAY--LEBTVDIADPPPIRQGEVNTVAVDGTFVLSCVATGSPFTILWRKDGVLV 495  
Qy 449 -----QLTNGHL---LSPJGG--GRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTS 496  
Db 496 STODSRIKQLENGVLQIRYAKLGDGTGYTCIASTPSGEA----- 534

[illegible]

Search completed: October 6, 2004, 18:20:40  
Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 12:45:02 ; Search time 88 Seconds  
(without alignments)  
2886.480 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVEPGLWPALLGIVLAWL.....AVAGLGQPDAGLFTVSBAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4787	100.0	899	5	AAU79939 Human UNC
2	4698.5	98.2	898	5	AAU85403 Human pro
3	4688.5	97.9	898	5	AAU97899 Human net
4	4545.5	95.0	898	2	AAW78898 Rat UNC-5
5	4545.5	95.0	898	5	AAU10543
6	4545.5	95.0	898	5	AAU97900
7	4434	92.6	943	4	AAU79128 Human pro
8	4379.5	91.5	842	5	AAU74818 Human REP
9	2789	58.3	556	2	AAW78899 Human UNC
10	2673.5	55.8	931	7	AB50691 Human UNC
11	2673.5	55.8	931	7	ADE63098 Human UNC
12	2673.5	55.8	932	4	ABG11551
13	2497	52.2	945	7	ADE63096 Rat Prote
14	2490.5	52.0	946	5	ABG61795 Novel UNC
15	2490.5	52.0	943	2	AAW78900 Rat UNC-5
16	2484	51.9	933	5	AAO18734 Human NOV
17	2479	51.8	945	4	AAU12244 Human NOV
18	2479	51.8	945	6	ABO17688 Human PRO
19	2479	51.8	945	6	ABU80942 Human PRO
20	2479	51.8	945	6	ABU66642 Human PRO
21	2479	51.8	945	6	ABU59723 Novel sec
22	2479	51.8	945	6	ABO24913 Human sec
23	2479	51.8	945	6	ABU66918 Human sec
24	2479	51.8	945	6	ADA45665 Novel hum
25	2479	51.8	945	6	ADA45665

## ALIGNMENTS

## RESULT 1

AAU79939  
ID AAU79939 standard; protein; 899 AA.

XX AC AAU79939;

XX AC (first entry)

DT 15-JUL-2002

XX Human UNC5-like protein NOV1.

DE Human UNC5-like protein NOV1.

XX Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;

KW cell signal processing; metabolic pathway modulation; cancerous tissue;

KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;

KW chromosome 13.

XX Homo sapiens.

OS WO200229038-A2.

XX 11-APR-2002.

PD 04-OCT-2001; 2001WO-US031377.

XX 04-OCT-2000; 2000US-0237862P.

XX (CURA-) CURAGEN CORP.

XX Herrmann JL, Rastelli L, Shinkets RA;

XX WPI; 2002-340104/37.

XX N-PSDB; ABK49422.

XX Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for

XX treating cardiomyopathy, atherosclerosis, and cancer.

XX Claim 1; Page 9; 180pp; English.

XX The present invention relates to a new NOVX polypeptide having a 900

XX (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)

XX residue amino acid sequence, as given in the specification. The novel

XX polypeptide, and its encoding polynucleotide, are used to treat

XX cardiomyopathy, atherosclerosis, cancer or a disease related to cell

XX signal processing and metabolic pathway modulation, in a human. Detecting

XX the polypeptide or polynucleotide is useful for identifying cancerous

XX tissue. The antibody can be used to treat diabetes or cancer. The host

XX cells can be used to produce non-human transgenic animals useful in drug

XX screening. The present amino acid sequence is that of the human UNC5-like

XX protein NOV1 of the invention. This sequence is encoded by the human UNC5

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC	-like NOV1 gene located on chromosome 13	
XX		
SQ	Sequence 899 AA;	
	Query Match 100.0%; Score 4787; DB 5; Length 899;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 899; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MAVRGLWPAALLGIVLAALRGSGAQAQSAIVANVPVPGANPDLLPHFIVPEPDVYIVNKP 60	
Db	1 MAVRGLWPAALLGIVLAALRGSGAQAQSAIVANVPVPGANPDLLPHFIVPEPDVYIVNKP 60	
Qy	61 VLLVCKAVPATQIFPKCNQGEVROVDHVIERSTGSGEPTMEVRIINVSQOQVEKVGLE 120	
Db	61 VLLVCKAVPATQIFPKCNQGEVROVDHVIERSTGSGEPTMEVRIINVSQOQVEKVGLE 120	
Qy	121 EYWCQCVAMSSGTTKQKAYIRIARLRKNEQEPFLAKEVSLQGIPLPCRPPIPPAE 180	
Db	121 EYWCQCVAMSSGTTKQKAYIRIARLRKNEQEPFLAKEVSLQGIPLPCRPPIPPAE 180	
Qy	181 VEWLRNEDLVDPSPDPNNYITREHSLVVRQARLADTANYTCVAKNIIVARRSASAAIVY 240	
Db	181 VEWLRNEDLVDPSPDPNNYITREHSLVVRQARLADTANYTCVAKNIIVARRSASAAIVY 240	
Qy	241 VNGWSTWTEWSVCSACGRGWQKRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSDG 300	
Db	241 VNGWSTWTEWSVCSACGRGWQKRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSDG 300	
Qy	301 SWSPSKWSACGLDCTHWRSECDPAPRNGGEECQGTDLTRNCTSLCVHVSASGPDV 360	
Db	301 SWSPSKWSACGLDCTHWRSECDPAPRNGGEECQGTDLTRNCTSLCVHVSASGPDV 360	
Qy	361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLDSVADSSIIITSGFPQVSIKPSKADNPHL 420	
Db	361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLDSVADSSIIITSGFPQVSIKPSKADNPHL 420	
Qy	421 LTIQPDISTTTTQGSICPRDQSPKPOLNGLHLLPLGGGRHTLHSSPTSBAEEFVS 480	
Db	421 LTIQPDISTTTTQGSICPRDQSPKPOLNGLHLLPLGGGRHTLHSSPTSBAEEFVS 480	
Qy	481 RLSTQNTVFRSLPRGTSNNMTYGFNPLGRLMIPNTGISLLIPDPAIPRGKIYEIYTLHK 540	
Db	481 RLSTQNTVFRSLPRGTSNNMTYGFNPLGRLMIPNTGISLLIPDPAIPRGKIYEIYTLHK 540	
Qy	541 PEDVRLPLAGCQTLSPITVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLLKQSCGSGW 600	
Db	541 PEDVRLPLAGCQTLSPITVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLLKQSCGSGW 600	
Qy	601 EODVLHLGEEAPSHLYYQLEASACVFTQJGRFALVGEALSVAAAKRLKLLLPAPVAC 660	
Db	601 EODVLHLGEEAPSHLYYQLEASACVFTQJGRFALVGEALSVAAAKRLKLLLPAPVAC 660	
Qy	661 TSLEYNIRVYCLHDTHDALKEVQLEKQGLGQLIQEPVRLHFKDSYHNLRLSIHDPVSSL 720	
Db	661 TSLEYNIRVYCLHDTHDALKEVQLEKQGLGQLIQEPVRLHFKDSYHNLRLSIHDPVSSL 720	
Qy	721 WSKLLVSQYQIEPFVHIWNGTQRYLHCTFTLERSVSPSTSDIACKLWVQVEGDGQSFIN 780	
Db	721 WSKLLVSQYQIEPFVHIWNGTQRYLHCTFTLERSVSPSTSDIACKLWVQVEGDGQSFIN 780	
Qy	781 FNITKDTFAELLALLESAGVPALVGPFAKIPFLIROKIISLDDPPCRRGADWMTLQAK 840	
Db	781 FNITKDTFAELLALLESAGVPALVGPFAKIPFLIROKIISLDDPPCRRGADWMTLQAK 840	
Qy	841 LHLDSLHFFASKPSFTAMILNLEWHPFNGNLSQLAAAVAGLQOPDAGLFTVSEAC 899	
Db	841 LHLDSLHFFASKPSFTAMILNLEWHPFNGNLSQLAAAVAGLQOPDAGLFTVSEAC 899	
RESULT 2		
AAU85403		
ID	AAU85403 standard; protein; 898 AA.	
XX		

AC	AAU85403;	
XX		
DT	21-MAY-2002 (first entry)	
XX		
DE	Human protein NOV1.	
XX		
DE	Human; NOVX; cardiomyopathy; atherosclerosis; diabetes;	
XX	cell signal processing disorder; metabolic disorder; obesity; infection;	
KW	anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;	
KW	Alzheimer's disease; Parkinson's disease; pain; asthma; hypertension;	
KW	haematopoietic disorders; dyslipidaemia; multiple sclerosis; angina pectoris;	
KW	osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;	
KW	myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;	
KW	psychosis; neurological disorder; anxiety; schizophrenia;	
KW	manic depression; dementia; dyskinesia; Huntington's disease;	
KW	Gilles de la Tourette's syndrome; gene therapy.	
XX		
XX	Homo sapiens.	
OS		
XX	WO200210216-A2.	
XX		
XX	07-FEB-2002.	
FD		
XX		
PF	30-JUL-2001; 2001WO-US024225.	
XX		
XX	28-JUL-2000; 2000US-0221409P.	
PR	04-AUG-2000; 2000US-0222840P.	
PR	04-AUG-2000; 2000US-0223752P.	
PR	04-AUG-2000; 2000US-0223762P.	
PR	04-AUG-2000; 2000US-0223769P.	
PR	14-AUG-2000; 2000US-0223770P.	
PR	14-AUG-2000; 2000US-0225146P.	
PR	15-AUG-2000; 2000US-0225392P.	
PR	15-AUG-2000; 2000US-0225470P.	
PR	16-AUG-2000; 2000US-0225697P.	
PR	01-FEB-2001; 2001US-0263662P.	
PR	05-APR-2001; 2001US-0281645P.	
XX		
XX	(CURA-) CURAGEN CORP.	
PI	Padigaru M, Mezès P, Mishra V, Burgess C, Casman S, Grose WM;	
PI	Alsebrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;	
XX		
XX	WPI; 2002-180074/23.	
DR	N-PSDB; ABK37922.	
XX		
PT	New isolated cytoplasmic, nuclear, membrane bound, or secreted	
PT	polypeptide, useful for treating cardiomyopathy, atherosclerosis,	
PT	infections, cancer, neurodegenerative, metabolic, hematopoietic and	
PT	immune disorders.	
XX		
PS	Claim 1; Page 11; 213pp; English.	
XX		
CC	The invention relates to an isolated cytoplasmic, nuclear, membrane	
CC	bound, or secreted polypeptide (NOVX, x=1-14) their variants or mature	
CC	form. Also included are the nucleic acids encoding the NOVX proteins, a	
CC	vector comprising the nucleic acid, a cell comprising the vector, an anti	
CC	-NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the	
CC	antibody are useful for treating or preventing a NOVX-associated	
CC	disorder, where the disorder is selected from cardiomyopathy,	
CC	atherosclerosis, diabetes, a disorder related to cell signal processing	
CC	and metabolic pathway modulation, metabolic disorders, obesity,	
CC	infectious disease, anorexia, cancer-associated cachexia, cancer,	
CC	neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,	
CC	immune disorders, hematopoietic disorders, and the various	
CC	dyslipidaemias, metabolic disturbances associated with obesity, the	
CC	metabolic syndrome X and wasting disorders associated with chronic	
CC	diseases, bacterial, fungal, protozoal and viral infections, pain,	
CC	bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's	
CC	disease, multiple sclerosis, Albright Hereditary Osteopetrosis, angina	
CC	pectoris, myocardial infarction, ulcer, allergy, benign prostatic	
CC	hypertrophy, and psychotic and neurological disorders, including anxiety,	
CC	schizophrenia, manic depression, delirium, dementia, and dyskinesias,	



Best Local Similarity 98.4%; Pred. No. 0;		Matches 886; Conservative 3; Mismatches 8; Indels 3; Gaps 3;	
QY	1	MAVRPGLWPALLGIVLAANLWGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVIVKKNP	60
DB	1	MAVRPGLWPALLGIVLAANLWGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVIVKKNP	60
QY	61	VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQVEKVFGL	120
DB	61	VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQVEKVFGL	120
QY	121	EYWCQCVAWSSGTTKSKQYIRIARLRKNEQEPQPLAKEVSLQGIIVLPCRPPEGIPPAE	180
DB	121	EYWCQCVAWSSGTTKSKQYIRIARLRKNEQEPQPLAKEVSLQGIIVLPCRPPEGIPPAE	180
QY	181	VEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSASAAVIV	240
DB	181	VEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSASAAVIV	240
QY	241	VNGGWSTWTSVCSACGRGWQRKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG	300
DB	241	VNGGWSTWTSVCSACGRGWQRKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG	299
QY	301	SWSPWKSACGLDCTHWRSECDPAPRNGEGCQGTDLDRCTSDLCVHSASGPEDEV	360
DB	300	SWSPWKSACGLDCTHWRSECDPAPRNGEGCQGTDLDRCTSDLCVHSASGPEDEV	359
QY	361	ALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL	420
DB	360	ALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDVADSSILTSFGFQVSIKPSKADNPHL	419
QY	421	LTIQPDLS-TTTTVQGLSPQDQSPKQFQTLNTHLLSLPLGGHRLTHSSPTSEAEFV	479
DB	420	LTIQPDLS-TTTTVQGLSPQDQSPKQFQTLNTHLLSLPLGGHRLTHSSPTSEAEFV	479
QY	480	SRLSTQNYFRSLPRTSNMTYGTNFIAGRLMIPNTGISLIPDPAIPRGKIYEIYTLH	539
DB	480	SRLSTQNYFRSLPRTSNMTYGTNFIAGRLMIPNTGISLIPDPAIPRGKIYEIYTLH	539
QY	540	KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPDPSLSLKKQSCGSG	599
DB	540	KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPDPSLSLKKQSCGSG	599
QY	600	WEQVILHGEAPSHLYCOLEASACVYFTQLGRFALVGEALSVAARLKLILFAPVA	659
DB	600	WE-DVLHGEAPSHLYCOLEASACVYFTQLGRFALVGEALSVAARLKLILFAPVA	658
QY	660	CTSEYNIRVYCLDTHDALKEVQLEKQGLQIEPRVLFHFKDSYHNLSIHDPVSS	719
DB	659	CTSEYNIRVYCLDTHDALKEVQLEKQGLQIEPRVLFHFKDSYHNLSIHDPVSS	718
QY	720	LWKSLLVSYQEIPIFYHWNQTRYLHCTFTLERSVSPSTSLACKLWVQVEGDQSPSI	779
DB	719	LWKSLLVSYQEIPIFYHWNQTRYLHCTFTLERSVSPSTSLACKLWVQVEGDQSPSI	778
QY	780	NFNITKTRFELLALAESEAGVPALVGSAPKIPFLIRQKLISSLDPPCRGADWRTLAQ	839
DB	779	NFNITKTRFELLALAESEAGVPALVGSAPKIPFLIRQKLISSLDPPCRGADWRTLAQ	838
QY	840	KLHLDLSLSPFASKPSPTAMILNWEARHPFNGLSQAALAAVAGLQPDAGLFTVSEAE	899
DB	839	KLHLDLSLSPFASKPSPTAMILNWEARHPFNGLSQAALAAVAGLQPDAGLFTVSEAE	898

RESULT 4

AAW78998

ID AAW78998 standard; protein; 898 AA.

XX AAW78998;

AC AAW78998;

XX 25-MAR-2003 (revised)

DT 21-DEC-1998 (first entry)

XX

DE	Rat UNC-5 homologue UNC5H-1.
XX	UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;
KW	diagnosis; therapy.
XX	Rattus sp.
XX	Key Location/Qualifiers
FT	Peptide 580..594
FT	/note="peptide used to raise rabbit polyclonal antisera"
XX	WO9837085-A1.
XX	27-AUG-1998.
XX	19-FEB-1998; 98WO-US003143.
XX	19-FEB-1997; 97US-00808982.
XX	(REGC ) UNIV CALIFORNIA.
XX	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;
XX	WPI; 1998-495364/42.
XX	N-PSDB; AAV52940.
XX	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and
XX	the biopharmaceutical industry.
XX	Claim 1; Page 19-22; 32pp; English.
XX	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis
XX	elegans UNC-5 protein. Their amino acid sequences were deduced from
XX	isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an
XX	E18 brain cDNA library. The predicted proteins show similarity with UNC-
XX	5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin
XX	type-1 repeats, a predicted membrane spanning region, and a large
XX	intracellular domain. They are predicted to be involved in cell migration
XX	and axon guidance, and are characterised as receptor proteins for
XX	netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins
XX	are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly
XX	from transfected host cells. The invention also provides unc-5
XX	hybridisation probes and primers, vertebrate UNC-5-specific binding
XX	agents such as specific antibodies, and methods of making and using the
XX	subject compositions in diagnosis (e.g. genetic hybridisation screens for
XX	vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate
XX	vertebrate unc-5 gene expression) and in the biopharmaceutical industry
XX	(e.g. as immunogens, reagents for modulating cell guidance, reagents for
XX	screening chemical libraries for lead pharmacological agents, etc.).
XX	(Updated on 25-MAR-2003 to correct PI field.)
XX	Sequence 898 AA;
QY	Query Match 95.0%; Score 4545.5; DB 2; Length 898;
DB	Best Local Similarity 94.7%; Pred. NO. 0;
QY	Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
DB	1 MAVRPGLWPALLGIVLAANLWGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVIVKKNP 60
QY	1 MAVRPGLWPALLGIVLAANLWGSAQOQSAIVANVPGANPDLLPHFLVEPEDEVIVKKNP 60
DB	61 VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQVEKVFGL 120
QY	61 VLLVCKAVPATQIFPKCNQGEWVRQVDHVIERTSDGSGEPTMEVRINVSQQQVEKVFGL 120
DB	121 EYWCQCVAWSSGTTKSKQYIRIARLRKNEQEPQPLAKEVSLQGIIVLPCRPPEGIPPAE 180
QY	121 EYWCQCVAWSSGTTKSKQYIRIARLRKNEQEPQPLAKEVSLQGIIVLPCRPPEGIPPAE 180
DB	181 VEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSASAAVIV 240
QY	181 VEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRRSASAAVIV 240

QY 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNVHRTVSSLLVSDG 300  
 Db |||||  
 QY 301 SWSPKWSACGLDCTHWSRECSDPAPRNGGECOGTDLDRNCTSDLCVHSASGPEDV 360  
 Db |||||  
 QY 361 ALYVGLIAVAVCLVLLLVLCYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420  
 Db |||||  
 QY 421 LTIQPDLS-TTTTYQGSICPRQDGPSPKQTLNTHLLSPLGGGRHTLHSSPTSEAEFV 479  
 Db |||||  
 QY 480 SRLSTONYFRSLPRGTSNMAYGTNFGRLMIPNTGISLLIPDAIPRGKIYEIYLTILH 539  
 Db |||||  
 QY 540 KPEDVRLPLAGCQTLISPIVSCGPGVLLTRPVILAMDHCGEPSPDSWSLRKKQSCGS 599  
 Db |||||  
 QY 600 WEQVHLHGEAPSHLYYCQLEASACYYVTEQLGRFALVGEALSVAALKRLKLLFAPVA 659  
 Db |||||  
 QY 660 CTSLEYNIRVYCLDTHDALKEVQLEKQGGQIQEPRVLHFKDSYHNLRSLSHDVPS 719  
 Db |||||  
 QY 720 LWSKLLVSYQEIPIFYHWNQYRLHCTFTLERVSPSTDLACKVWVQVGGQGSFSI 779  
 Db |||||  
 QY 780 NFNITKDTFAELIALSESGVPALVGFSAFKIFLIRQKIIISLDPPCRGADWRTLAQ 839  
 Db |||||  
 QY 840 KLHLDLSLFFASKPSATMILNWEARHPNGNLQALAAVAGLQPDAGLFTVSAEC 899  
 Db |||||  
 RESULT 5  
 AAU10543  
 ID AAU10543 standard; protein; 998 AA.  
 AC AAU10543;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Rat netrin receptor UNC5H1 (YSG7) polypeptide.  
 XX  
 KW YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;  
 KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;  
 KW calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;  
 KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;  
 KW tumour necrosis factor alpha; TNF-alpha; rat.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200175440-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 02-APR-2001; 2001WO-GB001486.  
 XX  
 PR 31-MAR-2000; 2000GB-00007880.  
 PR 26-MAY-2000; 2000GB-00012769.  
 XX  
 PW (WELF-) WELFIDE CORP.

XX  
 PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;  
 XX WPI; 2002-010813/01.  
 DR N-PSDB; AAS16843.  
 XX  
 PT Novel chronic animal model of schizophrenia, useful for identifying anti-  
 PT psychotic drugs and genes that are associated with schizophrenia.  
 XX  
 PS Disclosure; Fig 8b; 79pp; English.  
 XX  
 CC The invention relates to YSG polynucleotide fragments for use in  
 CC diagnosing and/or developing treatments for schizophrenia using chronic  
 CC animal models. The polynucleotides and their encoded polypeptides are  
 CC used for identification of compounds which modulate the expression of YSG  
 CC molecules, leading to the manufacture of schizophrenia medications. The  
 CC sequences can also be used for testing candidate compounds for any effect  
 CC on the polypeptides. Anti-schizophrenic effects of a compound can be  
 CC determined by measuring local cerebral glucose utilisation (LCGU) or  
 CC comparing its expression level with that of a control group. The  
 CC sequences are useful in the identification of genes associated with  
 CC schizophrenic states and in the development of an antibody. The sequences  
 CC of the invention include phosphodiesterase 1-alpha, calcium-independent  
 CC alpha-latrotoxin receptors (CIRL)-1,2&3 epithelial discoidin domain  
 CC receptor 1 (trkE), netrin receptor (UNC5H1), synapsins 1A and 1B and  
 CC tumour necrosis factor (TNF) alpha. This sequence represents rat netrin  
 CC receptor UNC5H1 (YSG7) polypeptide  
 XX  
 SQ Sequence 898 AA;  
 Query Match 95.0%; Score 4545.5; DB 5; Length 898;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;  
 QY 1 MAVRPGLPALLGLVLAALRGSGAQOSATVPANPGANPDLLPHLEVEDVIVKNKP 60  
 Db |||||  
 QY 61 VLLVCKAVPATQIFKCNGBWVRQVDHIERSTDGSSGSEPTMEVIRINVSROQVKVFGLE 120  
 Db |||||  
 QY 121 EYWCOCVAMSSGTTKQKAYIRIARLKNPQEPPLAKEVSLQGVILPCRPPIGPAE 180  
 Db |||||  
 QY 181 VEWLNRNEDLVDPNVIYITREHSLVVRQARLADTANYTCVAKNIVARRSAAVIVY 240  
 Db |||||  
 QY 241 VNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFCEGQNV-QKTACATLCEVDG 300  
 Db |||||  
 QY 301 SWSPKWSACGLDCTHWSRECSDPAPRNGGECOGTDLDRNCTSDLCVHSASGPEDV 360  
 Db |||||  
 QY 361 ALYVGLIAVAVCLVLLLVLCYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHL 420  
 Db |||||  
 QY 421 LTIQPDLS-TTTTYQGSICPRQDGPSPKQTLNTHLLSPLGGGRHTLHSSPTSEAEFV 479  
 Db |||||  
 QY 480 SRLSTONYFRSLPRGTSNMAYGTNFGRLMIPNTGISLLIPDAIPRGKIYEIYLTILH 539  
 Db |||||  
 QY 540 KPEDVRLPLAGCQTLISPIVSCGPGVLLTRPVILAMDHCGEPSPDSWSLRKKQSCGS 599  
 Db |||||

[illegible]

RES. IT. 6

RESULT: 6  
AAT97900

AAU97900  
ID AAU97900 standard: protein; 898 AA.

XX  
AC  
ATT07000.

XX  
DT 27-AUG-2002 (first entry)

XX  
DT 27-AUG-2002 (first entry)

XX DE Rat netrin binding membrane receptor UNC5H-1 protein.

Netrin binding membrane receptor; receptor; UNC5H-1; Rat; neurotropic;  
neuroprotective; cytostatic; antiparkinsonian; cerebroprotective; cancer;  
central nervous system; CNS; stroke; parkinson's disease;  
multiple sclerosis; Alzheimer's disease.

OS *Rattus* sp.

XX	Key	Location/Qualifiers
FH		

FT	Domain	152. .223	Location/Qualifiers
FT	key		

```

FT      domain
      /note= "Immunoglobulin domain "

```

FT Domain 247, .294

FT	name= "Thrombospondine type 1 domain
202	248
203	249

FT	Domain
FT	302. .348
FT	/note- "Thrombospandine type 1 domain"

FT	Region	361	.382
FT		/note="Thrombospandine type I domain	

```

FI REGION
361. .382
/note= "Transmembrane region"
FI

```

```

FT      Domain
FT      495.598
FT      /acc= transmembrane region

```

```
FT /note= "ZU5 domain"
```

FT	Domain	817.	.897

FT /note= "Death domain"

XX  
DN W0708933000-12

PN W0200233080-A2.  
XX

XX 25-APR-2002.  
PD

XX  
172  
172002-VIR-02

PF 15-OCT-2001; 2001WO-EP011891.

Query Match	95.0%;	Score 4545.5;	DB 5;	Length 898;
Best Local Similarity	94.7%;	Pred. No. 0;		
Matches 852.	Conservative	19;	Mismatches 26;	Indels 3;
				Gaps 3;



QY 720 LWSKLLVSQEIPIFYHWMGTORYLHCTFTLERSVSPSTSDLACKLWVWQVGGQSPSI 779  
 Db 719 LWSKLLVSQEIPIFYHWMGTQYLYLHCTFTLERINASTDLACKWVWQVGGQSFNI 778  
 QY 780 NFNTKTRFAELLALSEAGVPALVGFSAFKIPFLIRQKIISLDPFCRRGADWRTLAQ 839  
 Db 779 NFNTKTRFAELLALSEAGVPALVGFSAFKIPFLIRQKIISLDPFCRRGADWRTLAQ 838  
 QY 840 KLHLDLSLSPFASKPSTAMILNWEARHPFNGNLQALAAVAGLQDPDAGLFTVSEAC 899  
 Db 839 KLHLDLSLSPFASKPSTAMILNWEARHPFNGNLQALAAVAGLQDPDAGLFTVSEAC 898

RESULT 7  
 AAM79128  
 ID AAM79128 standard; protein; 943 AA.  
 XX AC AAM79128;  
 XX DT 06-NOV-2001 (first entry)  
 XX DE Human protein SEQ ID NO 1790.  
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX OS Homo sapiens.  
 XX PN WO200157190-A2.  
 XX PD 09-AUG-2001.  
 XX PF 05-FEB-2001; 2001WO-US004098.  
 XX PR 03-FEB-2000; 2000US-00496914.  
 PR 27-APR-2000; 2000US-00560875.  
 PR 20-JUN-2000; 2000US-00598075.  
 PR 19-JUL-2000; 2000US-00620325.  
 PR 01-SEP-2000; 2000US-00654936.  
 PR 15-SEP-2000; 2000US-00663561.  
 PR 20-OCT-2000; 2000US-00693325.  
 PR 30-NOV-2000; 2000US-00728422.  
 XX (HYSE-) HYSEQ INC.  
 XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX WPI; 2001-476283/51.  
 DR N-PSDB; AAK52261.  
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 in diagnosis and gene therapy.  
 Claim 20; Page 4148-4150; 6221pp; English.  
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to  
 cytokine, cell proliferation or cell differentiation or which may induce  
 production of other cytokines in other cell populations. The  
 polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 peptide therapy. The polypeptides have various cytokine-like activities,  
 e.g. stem cell growth factor activity, haematopoiesis regulating  
 activity, tissue growth factor activity, immunomodulatory activity and  
 activin/inhibin activity and may be useful in the diagnosis and/or  
 treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 sequence listing were missing at the time of publication

XX SQ Sequence 943 AA;  
 Query Match 92.6%; Score 4434; DB 4; Length 943;  
 Best Local Similarity 90.4%; Pred. No. 0;  
 Matches 853; Conservative 4; Mismatches 17; Indels 70; Gaps 7;  
 QY 1 MAYRPGIWPALLGIVLAALWLRGSAQOQSATVAVPVGANPDLLPHFLVEPDEVIVNKP 60  
 Db 25 MTRRPSL-----MAGRQHGSQAQOQSATVAVPVGANPDLLPHFLVEPDEVIVNKP 76  
 QY 61 VLVVCKAVPATQIFFKCNCEWVRQDVHVIERSDTGSGSEPTMEVRINVSQOQVKEVFGLE 120  
 Db 77 VLVVCKAVPATQIFFKCNCEWVRQDVHVIERSDTGSGSEPTMEVRINVSQOQVKEVFGLE 136  
 QY 121 EYWCQCVMSSSGTTKQKAYIRIARLRNFEQPLAKEVLSLEQGVLPQCRPEGIPPAE 180  
 Db 137 EYWCQCVMSSSGTTKQKAYIRIARLRNFEQPLAKEVLSLEQGVLPQCRPEGIPPAE 196  
 QY 181 VEMLRNEDLVDPISLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240  
 Db 197 VEMLRNEDLVDPISLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 256  
 QY 241 -----VNGCWSTWTSVCSASCGRCWOKRSRSTN 271  
 Db 257 GPRDSILVTGRGTAVPLGSDMWLSFVVRPVNGGWSTWTSVCSASCGRCWOKRSRSTN 316  
 QY 272 PAPLNGAFCEGQNVHDRTVSSLLVSDGSPWSKWSACGLDCTHWSRECSDDPAPRNG 331  
 Db 317 PAPLNGAFCEGQNV-KTACATLCPVDGSPWSKWSACGLDCTHWSRECSDDPAPRNG 375  
 QY 332 GEBCQGTDLDRNCTSDLCVH-----SASGPEVALYVGLIAVAVCLVLL 376  
 Db 376 GEBCQGTDLDRNCTSDLCVHNYTPTAKMLSPAAASGPEVALYVGLIAVAVCLVLL 435  
 QY 377 LVLVLVYCKRKEGLSDVDADSSILTSGFQPSIKPSKADNPHLLITQPDLS-TTTTYQG 435  
 Db 436 LVLVLVYCKRKEGLSDVDADSSILTSGFQPSIKPSKADNPHLLITQPDLS-TTTTYQG 495  
 QY 436 SLCPRODGPSPKQLTNGHLLSPGGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGT 495  
 Db 496 SLCPRODGPSPKQLTNGHLLSPGGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGT 555  
 QY 496 SNMTYGTFFLGGRLMPNTGISLIPPDAPRKGKIYEIYLTHKPEDVRLPAGCOTLL 555  
 Db 556 SNMTYGTFFLGGRLMPNTGISLIPPDAPRKGKIYEIYLTHKPEDVRLPAGCOTLL 603  
 QY 556 SPVSCGPPGVLLTRPVILAMDRHCGEPPSDSNLRLKQSCGSEWEDVHLGEEAPSHL 615  
 Db 604 ---VSCGPPGVLLTRPVILAMDRHCGEPPSDSNLRLKQSCGSEWEDVHLGEEAPSHL 659  
 QY 616 YYQLEASACYVTEQGRFALVGEALSAVAAKRLKLLFPAPVACTSLEYNIRVYCLHDT 675  
 Db 660 YYQLEASACYVTEQGRFALVGEALSAVAAKRLKLLFPAPVACTSLEYNIRVYCLHDT 719  
 QY 676 HDALKEVQLEKQGLQIQRVHLHFQDSYHNLRLSIHDVPSLLKSKLLSYQIPFY 735  
 Db 720 HDALKEVQLEKQGLQIQRVHLHFQDSYHNLRLSIHDVPSLLKSKLLSYQIPFY 779  
 QY 736 HIWNGTQRYLHCTFTLERSVSPSTDLACKLWVWQVGGQSPSINFNITKTRFAELLAL 795  
 Db 780 HIWNGTQRYLHCTFTLERSVSPSTDLACKLWVWQVGGQSPSINFNITKTRFAELLAL 839  
 QY 796 ESEAGVPALVGFSAFKIPFLIRQKIISLDPFCRRGADWRTLAQKLHLSLSPFASKPS 855  
 Db 840 ESEAGVPALVGFSAFKIPFLIRQKIISLDPFCRRGADWRTLAQKLHLSLSPFASKPS 899  
 QY 856 PTAMILNLWEARHPFNGNLQALAAVAGLQDPDAGLFTVSEAC 899  
 Db 900 PTAMILNLWEARHPFNGNLQALAAVAGLQDPDAGLFTVSEAC 943

AAU74818  
 ID AAU74818 standard; protein; 842 AA.  
 AC AAU74818;  
 XX  
 XX  
 DT 23-APR-2002 (first entry)  
 DE Human REPTR 1 protein.  
 XX  
 XX REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;  
 KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;  
 KW antiallergic; antibody; immunogen; endometriosis;  
 KW gastrointestinal disorder; gastritis; oesophageal carcinoma;  
 KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;  
 KW endocrine disorder; hypothalamus disorder; Kallman's disease;  
 KW autoimmune disease; inflammatory disease; infertility; receptor;  
 KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;  
 KW osteoarthritis; diabetes mellitus; multiple sclerosis;  
 KW systemic lupus erythematosus; cell proliferative disorder; cancer;  
 KW developmental disorder; Duchenne muscular dystrophy;  
 KW Becker muscular dystrophy; neurological disorder; epilepsy;  
 KW Alzheimer's disease; Huntington's disease; reproductive disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200198354-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 XX 21-JUN-2001; 2001WO-US019942.  
 XX  
 PR 21-JUN-2000; 2000US-0214027P.  
 PR 25-AUG-2000; 2000US-0228045P.  
 PR 12-DEC-2000; 2000US-0255104P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;  
 PI Lal P, Policky JL, Azimzai Y, Lu DAM, Graul R, Yao MG, Burford N;  
 PI Hafalia AJA, Baughn MK, Bandman O, Patterson C, Yang J, Xu Y;  
 PI Gandhi AB, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;  
 XX WPI: 2002-090432/12.  
 DR N-PSDB; ABK35169.  
 XX  
 XX Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in  
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
 PT gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell  
 PT proliferative (e.g. cancer) disorders.  
 XX  
 XX Claim 45; Page 111-113; 157pp; English.  
 XX  
 CC This invention relates to twelve human receptors cDNA sequences referred  
 CC to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The  
 CC proteins of the invention may have antiinflammatory, cytostatic,  
 CC immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active  
 CC general, anticonvulsant, nootropic, neuroprotective, antiallergic  
 CC activities. The sequences of the invention may be used to produce REPTR  
 CC agonists or antagonists, and the protein sequences may be used to raise  
 CC anti-REPTR antibodies. These molecules and the REPTR polynucleotides and  
 CC polypeptides of the invention are useful in the diagnosis, treatment and  
 CC prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,  
 CC Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine  
 CC (e.g. hypothalamus disorder, Kallman's disease), autoimmune/ inflammatory  
 CC (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,  
 CC allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,  
 CC systemic lupus erythematosus), cell proliferative (e.g. cancer),  
 CC developmental (e.g. Duchenne and Becker muscular dystrophy), neurological  
 CC (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and  
 CC reproductive (e.g. infertility, endometriosis) disorders. Numerous other  
 CC examples of each disorder are given in the specification. The present  
 CC sequence represents the human REPTR1 protein sequence of the invention  
 XX

SQ Sequence 842 AA;  
 Query Match 91.5%; Score 4379.5; DB 5; Length 842;  
 Best Local Similarity 93.1%; Pred. No. 0;  
 Matches 838; Conservative 1; Mismatches 2; Indels 59; Gaps 3;  
 QY 1 MAVRPGIWPALIGLIVLAALWLRGSGAQOSATVANPVPGANPDLLPHFLVEDEVDVIVKNKP 60  
 DB 1 MAVRPGIWPALIGLIVLAALWLRGSGAQOSATVANPVPGANPDLLPHFLVEDEVDVIVKNKP 60  
 QY 61 VLVKCAVPATQIFFKCNCEWVRQVDHVIERSDSSGSEPTMEVRINVSQQVEKVFGE 120  
 DB 61 VLVKCAVPATQIFFKCNCEWVRQVDHVIERSDSSGSEPTMEVRINVSQQVEKVFGE 120  
 QY 121 EYWCQCVAMSSSGTTKSQKAYIRIARLRKPEPEPLAKEVLSBOGIVLPCRPPEGIPPAE 180  
 DB 121 EYWCQCVAMSSSGTTKSQKAYIRIARLRKPEPEPLAKEVLSBOGIVLPCRPPEGIPPAE 180  
 QY 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240  
 DB 181 VEWLNRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVY 240  
 QY 241 VNGGWTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVHRTVSSLLVSDG 300  
 DB 241 -----VDG 243  
 QY 301 SWSPWKSWSACGLDCTHWSRECSDDPAPRNGBECQGTDLDRNCTSDLCVHSASGPEV 360  
 DB 244 SWSPWKSWSACGLDCTHWSRECSDDPAPRNGBECQGTDLDRNCTSDLCVHTASGPEV 303  
 QY 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDSDVADSSILTSGFQVPSKPSKADNPHL 420  
 DB 304 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDSDVADSSILTSGFQVPSKPSKADNPHL 363  
 QY 421 LTIQPDLS-TTTTYQGSCLCPRODQSPKFOITNGHLLSPLGGGRHTLHSSPTSEAEFFV 479  
 DB 364 LTIQPDLS-TTTTYQGSCLCPRODQSPKFOITNGHLLSPLGGGRHTLHSSPTSEAEFFV 423  
 QY 480 SRLSTQNYPSRLPRGTSNMTYTFNFGRLMIPNTGISLILIPDAIPRGKIYIYLT 539  
 DB 424 SRLSTQNYPSRLPRGTSNMTYTFNFGRLMIPNTGISLILIPDAIPRGKIYIYLT 483  
 QY 540 KPEDVRLPLAGCOTLLSPIVSCGPPGVLLTRPVILAMDHGCEPSPDSWSLRLKKQSCGGS 599  
 DB 484 KPEDVRLPLAGCOTLLSPIVSCGPPGVLLTRPVILAMDHGCEPSPDSWSLRLKKQSCGGS 543  
 QY 600 WEQDVLHLGBEAPSHLYYCOLEASACVYFTQGRFALVGEALSVAALKRLKLLFAPVA 659  
 DB 544 WE-DVLHLGBEAPSHLYYCOLEASACVYFTQGRFALVGEALSVAALKRLKLLFAPVA 602  
 QY 660 CTSLEYNIRVYCLDHTDALKVVOLEKQGLQIQEPVILHFKDSYHNRLSLTHDVPSS 719  
 DB 603 CTSLEYNIRVYCLDHTDALKVVOLEKQGLQIQEPVILHFKDSYHNRLSLTHDVPSS 662  
 QY 720 LMKSKLLVSVQEPFFYHINWGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGQGSFSI 779  
 DB 663 LMKSKLLVSVQEPFFYHINWGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGQGSFSI 722  
 QY 780 NFNTKDTREFAELALESEAGVPALVGPSAFKIPFLIRQKIISLSDPPCRGADWRTLAQ 839  
 DB 723 NFNTKDTREFAELALESEAGVPALVGPSAFKIPFLIRQKIISLSDPPCRGADWRTLAQ 782  
 QY 840 KLHLDSHLSFASKPSPMTAMLNWEARHPNGLSOLAAVAGLGOPDAGLFTVSEAC 899  
 DB 783 KLHLDSHLSFASKPSPMTAMLNWEARHPNGLSOLAAVAGLGOPDAGLFTVSEAC 842  
 RESULT 9  
 AAU78899  
 ID AAU78899 standard; protein; 556 AA.  
 XX  
 XX AC AAU78899;  
 XX

DT 25-MAR-2003 (revised)  
 DT 21-DEC-1998 (first entry)  
 DE Human UNC-5 homologue UNC5H-1.  
 XX  
 KW UNC-5; UNC5H-1; human; netrin receptor; cell migration; axon guidance;  
 KW diagnosis; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 7 /note= "encoded by TG"  
 FT Misc-difference 67 /note= "encoded by TG"  
 FT Misc-difference 256 /note= "encoded by AtCT"  
 FT Misc-difference 262 /note= "encoded by GC"  
 FT Misc-difference 360 /note= "encoded by TG"  
 FT Misc-difference 367 /note= "encoded by AG"  
 FT Misc-difference 370 /note= "encoded by CC"  
 FT Misc-difference 542 /note= "encoded by TC"  
 FT Misc-difference 542 /note= "encoded by GG"  
 PN WO9837085-A1.  
 XX  
 XX 27-AUG-1998.  
 XX  
 XX 19-FEB-1998; 98WO-US003143.  
 XX  
 XX 19-FEB-1997; 97US-00808982.  
 XX (REGC ) UNIV CALIFORNIA.  
 XX  
 XX Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;  
 XX WPI: 1998-495364/42.  
 DR N-PSDB; AAW78899.  
 XX  
 XX Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and  
 XX the biopharmaceutical industry.  
 XX  
 XX Claim 1; Page 22-23; 32pp; English.  
 XX  
 CC UNC5H-1 and UNC5H-2 (see AAW78901) are human homologues of Caenorhabditis  
 CC elegans UNC-5 protein. Their amino acid sequences were deduced from  
 CC isolated unc5h cDNA clones (see AAV52941 and AAV52943) isolated from an  
 CC embryonic brain cDNA library. The predicted proteins show similarity with  
 CC UNC-5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin  
 CC type-1 repeats, a predicted membrane spanning region, and a large  
 CC intracellular domain. They are predicted to be involved in cell migration  
 CC and axon guidance, and are characterised as receptor proteins for  
 CC netrins. Rat UNC5H-1 (see AAW78998) and UNC5H-2 (see AAW78900) proteins  
 CC are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly  
 CC from transfected host cells. The invention also provides unc-5  
 CC hybridisation probes and primers, vertebrate UNC-5-specific binding  
 CC agents such as specific antibodies, and methods of making and using the  
 CC subject compositions in diagnosis (e.g. Genetic hybridisation screens for  
 CC vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate  
 CC vertebrate unc-5 gene expression) and in the biopharmaceutical industry  
 CC (e.g. as immunogens, reagents for modulating cell guidance, reagents for  
 CC screening chemical libraries for lead pharmacological agents, etc.).  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 556 AA;

Query Match 58.3%; Score 2789; DB 2; Length 556;  
 Best Local Similarity 96.6%; Pred. No. 4.5e-233;  
 Matches 538; Conservative 1; Mismatches 16; Indels 2; Gaps 2;

QY 344 NCTSDLCVHSASGPDVALYVGLIAVAVCLVLLLVILVYCRKKEGLSDVADSSILTS 403  
 DB 1 NCTSDLCVHTASGPDVALYVGLIAVAVCLVLLLVILVYCRKKEGLSDVADSSILTS 60  
 QY 404 GFQPVSIKPSKADNPHELLTIQPDLS-TTTTVOGSLCPRODGPSPKFQLTNGHLLSPLOGG 462  
 DB 61 GFQPVSIKPSKADNPHELLTIQPDLS-TTTTVOGSLCPRODGPSPKFQLTNGHLLSPLOGG 120  
 QY 463 RHTLHSSPTSEAEFEVSRSLSTQNYFRSLPRGTSNMVTGTNFGRLMIPNTGISLLIP 522  
 DB 121 RHTLHSSPTSEAEFEVSRSLSTQNYFRSLPRGTSNMVTGTNFGRLMIPNTGISLLIP 180  
 QY 523 PDAIPRGKIYIYITLHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP 582  
 DB 181 PDAIPRGKIYIYITLHKPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEP 240  
 QY 583 SPDSWSLRLKKQSCGSEGWEDVLHGEAPSHLYVCOLEASACYVFTQGRFALVGEAL 642  
 DB 241 SPDSWSLRLKKQSCGSEGWEDVLHGEAPSHLYVCOLEASACYVFTQGRFALVGEAL 299  
 QY 643 SVAAAKRLKLLLPAPVACTSLEYNIRVYVCLDTHDALKKEVVOLEKQGGQLIQEPVRLHF 702  
 DB 300 SVAAAKRLKLLLPAPVACTSLEYNIRVYVCLDTHDALKKEVVOLEKQGGQLIQEPVRLHF 359  
 QY 703 KDSYHNLRSLSDHVPSSLWKSLLVSYQEIPIFYHNGTORYLHCTFTFLERSVSPSTSLA 762  
 DB 360 KDSYHNLRSLSDHVPSSLWKSLLVSYQEIPIFYHNGTORYLHCTFTFLERSVSPSTSLA 419  
 QY 763 CKLWVQVEGDGQSFNFINITKOTFAELLALSEAGVPALVGPSPAFKIPFLIROKIIIS 822  
 DB 420 CKLWVQVEGDGQSFNFINITKOTFAELLALSEAGVPALVGPSPAFKIPFLIROKIIIS 479  
 QY 823 SLDPCCRRGADWRTLAKLHLDLSLSPFASKPSTAMILNLWEARHPNGNLSOLAAAVA 882  
 DB 480 SLDPCCRRGADWRTLAKLHLDLSLSPFASKPSTAMILNLWEARHPNGNLSOLAAAVA 539  
 QY 883 GLGQPDAGLFTVSEAC 899  
 DB 540 GTPAGRLLSQSCSEAC 556  
 RESULT 10  
 AAB50691  
 ID AAB50691 standard; protein; 931 AA.  
 XX  
 AC AAB50691;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE Human UNC5C protein SEQ ID NO:90.  
 XX  
 KW Human; Caenorhabditis elegans; UNC-5; splice variant; nematode worm;  
 KW protein-protein interaction; identification.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000073328-A2.  
 XX  
 PD 07-DEC-2000.  
 XX  
 PF 02-JUN-2000; 2000WO-EP005108.  
 XX  
 PR 01-JUN-1999; 99GB-00012755.  
 XX (DEVG-) DEVGEN NV.  
 XX  
 PI Van Crieckinge W, Roelens I, Bogaert T, Verwaerde P;  
 XX WPI: 2001-016508/02.  
 XX  
 PT Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and a  
 PT human unc-5Hs1 cDNA, useful in yeast two hybrid experiments for

PT identifying unknown human cDNAs which encode proteins that interact with  
PT the human unc-5C protein.

PT the human unc-5C protein.

PS Disclosure: Page 224-227; 246pp; English.

The present invention describes 3 variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-5C transcripts, and a human unc-5Hs1 cDNA which shares homology with the *Rattus norvegicus* unc-5Hs1 cDNA. Also described are assays based on protein-protein-interactions between the unc-5 protein and a variety of different protein-interacting proteins. The unc-5C variant cDNAs and unc-5Hs1 cDNA are useful in methods for identifying compounds which reduce or inhibit the lethal phenotype associated with the expression of the unc-5 death domain in yeast. They are also useful in yeast two hybrid experiments for identifying unknown human cDNAs which encode proteins that interact with the human unc-5C protein. AAC90914 to AAC90971 and AAB50646 to AAB50693 represent sequences used in the exemplification of the present invention.

Sequence 931 AA;

Query Match

Query Match

Best Local Similarity 55.5%; Pred. No. 1.1e-222;  
Matches 507; Conservative 154; Mismatches 221;  
Indels 31; Gaps 12;

9 PALLGIVLAAWLRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVC 65

[illegible]

66 KAVPATOTFFKNCGEWVROVDHVIERSTDGSSGEPTMEVRINVSROOVEKVFGLEEYWCQ 125

[illegible]

DD 84 NASFAIQTITKONSEWVHQWJHTVDEKVDZTPOGFIANKVQVITITONXX,EEETOCDEEANCX

037 126 CVAWSSSGTTKSOKAYIRIARI.RKNFEOEPIAKEVSLEOGIVLPCRPEGIPPAEVEVWLR 185

144 CVAWSSAGTttKSRKAYVRIAYTRKTFEEOEPIGKEVSTEOEVTIOCRPPEGI PVAEVEWLK 203

DD I44 CVHWSAGIINBKNHI VNIATIAI UANITIEQEFEJGAKVUUEZYEVZDQXCKRLELOII VMEVPHNEAC 2003

QY 186 NEDLVDPSENVYITIREHSTLVVRQAKLADIANIICVAKNIIVARRKASCAAAVIVIVAGGN 2413

204 NEDTIDPVEDRNEFYTTIDHNLIIKOARLSDTANYTCVAKNIIVAKRKS TTATVIVVNGGW 263

246 STWTFWSVCASCCGPGWOKRSRSCITNPAPI.NCGAFCEGONTVHDR TVSS.I.VSVDGGSWPW 305

[illegible]

DD 204 SIMIENSVCNKRCKRGIVQARIKICINEFELNGHATCEQSV-QTACCTICEFADGANTH 022

306 SKWSACGIDCTHWRSPECSDPAPRNGGEECOGTDIDTRNCTSDLCVHSASGPEDVALYVG 365

323 SKWSTCGTECTHWRPRRRCCTAPAKNGGKDCNGI.VTOSKNCTGTGTCMOTAPSDPVALYVG 382

DB 323 SNW5ICGIECIHWRKRBCIAHFAFNWGNDCGFLVUQ3XNCLPQTCMQLHEDDDVALI19300Z

QY 366 L-LAVAVCLVLELVLLVICRKEGLDSDVADSSILFISQFVSTKFSKADNEHTITQ 424

383 TVTAVTVCTATSVVVAIFVYRKNNHRDFESDIIDSSALNGGFOPVNIKAAROD---LLAVP 439

425 PDI - STTTTYYVCGSI CPPBQNGPSPKFOI.TNGHILSPILGGGRHTI.HHS-----SPTSEAEFFV 479

QY 423 PDL-S11111Q8SLCFCRQD97SEKTQINGHDESEFEGGKATLHHS OF ICEBERG, 1972

DB 440 PDLTSAAMYRGVPYALHD-VSDKIPMINSFILDDPLFNLRKRVINISGHAVSFQDDESEFI 438

480 SPT.S-----TONVE-----RST.PRGT--SNMTYGTENFI.GGRI.MTPNTGISLLIPPDA 525

499 SVI SPOMTOSI I ENFAI SI KNOSI AROTDPSC TAE GSEFNSI GGHI I VPNSGVSI I TPAGA 558

DB 499 SNLSFQMIQSLLENEALS LKNNQSDAKQYIDFSCIAFGSFGNLSLOCHLIVFNCSVSLILFACA 5500

CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 931 AA;

Query Match 55.8%; Score 2673.5; DB 7; Length 931;  
 Best Local Similarity 55.5%; Pred. No. 1.1e-222;  
 Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;

QY 9 PALGIVLAWLRSQAQSS---ATVANPVGANPDLLPHFLVPEPDYVIVKNKPVLLVC 65  
 Db 26 PAL--ALLSASGTSAQDDFFHELPEPTSPDPPEPLHELPEEAYIVKNKPVNLYC 83  
 QY 66 KAVPATQIFKCGEWRQVDHVIERSDGSSEFTMEVRINVRQVEKVGLEETWCQ 125  
 Db 84 KASPATQIFYKCGEWRHQDHIIVDERVDETSGLIIVREVSIEISRQVVEELFGPEYWCQ 143  
 QY 126 CVAWSSGTTKSKAYIRIARLRKNFEQPLAKEVSLQGIPLCRPPEGIPPAEVEWLR 185  
 Db 144 CVAWSSAGTTKSKAYIRIARLRKTFEQLPKEVSLQGIPLCRPPEGIPPAEVEWLR 203  
 QY 186 NEDLVPSLPNVITREHSLVVRQALADPANTYCVAKNIVARRRSASAAVIVYVNGW 245  
 Db 204 NEDLIDVEDRNFYITIDHNLIIKQARLSDTANTYCVAKNIVAKRKSTTATVIVYVNGW 263  
 QY 246 STTWSVCSASCGRGWQKSRCTNPAPLNGAFCEGQNVHDTVSSLLVSDGSPW 305  
 Db 264 STTWSVCSNRGQYQKTRCTNPAPLNGAFCEGQSV-QKIACTTCLPVDGRWTPW 322  
 QY 306 SKWSACGLDCTHWRSCSDPAPNGGEECGQTDLTNRCTSDLCVHSASGPEVVALYVG 365  
 Db 323 SKWSTCGTECHWRRECTAPAPNGGKDCDGLVQSKNCTDGLCMQTAPDSDVVALYVG 382  
 QY 366 L-TAVAVCLVLLLVILVYCRKKGDSVDVASSILTSFGPQVSIKPSKADNPHELLTIQ 424  
 Db 383 IVIAVIVCLAVVVALFVYKKNRHPDESIIIDSSALNGGFQPNVIAKARQD---LLAVP 439  
 QY 425 PDL-STTTTYGSLCPQDGPSPKQFQLTNGHLLSPGGRHLLHS-----SPTSEAEFV 479  
 Db 440 PDLTSAAMTRGPVYALHD-VSDKIPMTNSPILDLPLNPKIKVNTSGAVSPQDDLSEFT 498  
 QY 480 SRLS---TQNYF-----RSLPRTG--SNMTYGTFFLGGRLMTPNTGISILLIPDA 525  
 Db 499 SKLSPQMTQSLLENEALS LKXNSLARQTDPSCTAFGSFNLGHLIVPNSGVSLILIPAGA 558  
 QY 526 IPGKYEIVLTHKPDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEFPD 585  
 Db 559 IPQGRVYEMVTYVREKTRPMPDSDQTLTLTPVVVSCGPGALLTRPVILTMHCADNTE 618  
 QY 586 SWSLRLKQKSCGSEWDLHLGEEAPSHLYYQLEASACVYVTEQGRPALVGEALSVA 645  
 Db 619 DWKILLKQNAQGWGE-DVVVWGEENFTTFCYIKLDAEACHILTNLSLVALVGHSTTKA 677  
 QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKVEVVQLEKOLGGOLIOEPRLVHKDS 705  
 Db 678 AAKRLKLAIFGLPCCSLEYSIRVYCLDTHDALKVEVVQLEKOLGGOLIOEPRLVHKDS 737  
 QY 706 YHNLRLSHDVPSSLKSLVSYQEIPIFYHWNQYRVLHCTFTLERSVPSSTDLACKL 765  
 Db 738 THNLRSLIHDIAHSLKSLVSYQEIPIFYHWNQYRVLHCTFTLERSVPSSTDLACKL 797  
 QY 766 WYMQVEGDSGFSINFTYKDFRFAELLESAGVPALVGPFAKIPPIRQKILSSLD 825  
 Db 798 CVRQVEGEGQIFQINCTVSEPTGIDLPDLPDANTITVTGSPAFSIPILRQKILSSLD 857

QY 826 PPCRGADWRTLAOKLHLDLSLSPFASKPSPPTAMILNLEARHPNGNLSQLAARAVAGLG 885  
 Db 858 APQTRGHDWRMLAKHNLDRYLNIFATKSSPTGVILDLWEAQNPDPGNLSMLAAVLEEMG 917  
 QY 886 QPDAGLFTVSEAE 898  
 Db 918 RHEIVSVSLAEEGQ 930

RESULT 12

ABG11551  
 ID ABG11551 standard; protein; 982 AA.

XX AC ABG11551;  
 XX DT 18-FEB-2002 (first entry)  
 XX DE Novel human diagnostic protein #11542.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS75738.

XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX Claim 20; SEQ ID NO 41910; 103bp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 982 AA;

Query Match 55.8%; Score 2673.5; DB 4; Length 982;  
 Best Local Similarity 55.5%; Pred. No. 1.3e-222;

Matches	507;	Conservative	154;	Mismatches	221;	Indels	31;	Gaps	12;					
QY	9	PALLGIVLA	AWLRGSAQGS	---	ATVANPVG	ANPDLL	PHPLVPE	PVYIVKKNPVL	LVC 65					
Db	77	PAL	--ALLSAGT	SAQAQDD	FFHEL	PETPSP	DPPEPL	PHPLIEPE	AYIVKKNPVNLYC 134					
QY	66	KAVPATOI	FFKNCGEW	RVQDVH	VIERT	SDSSG	SEPTME	VRINVSR	QOVKEVFGLEEYWCQ 125					
Db	135	KASPATOI	FFKNCSEW	HQDHIV	ERDETS	GLIVREV	SEI	SRQVVEEL	FGPEYWCQ 194					
QY	126	CVAMSSG	TTKSKAYI	RIARL	RNFQEP	ELAKEV	SLQEGIV	LPFCRP	PEGIPPAEVEWLR 185					
Db	195	CVAMSSA	GTTKSKAYI	VIAYLR	KTRTFE	QPLGKEV	SLQEVLL	QCRPE	GIPPAEVEWLK 254					
QY	186	NEDLVPS	LDNPVYIT	REHSLV	VRQARL	ADTANYT	CVAKNIV	ARRRSASA	AVIYVNGW 245					
Db	255	NEDIIDP	VEDRNF	YITDHN	LIIRKQ	ARLSDTANYT	CVAKNIV	ARRRSASA	AVIYVNGW 314					
QY	246	STWTEW	SVCSAS	CGRGWK	RSRSC	TNPAPL	NGGAF	CEGONV	HDTRVSSLIVSVDGNS	WP 305				
Db	315	STWTEW	SVCSNR	CGRGYQ	KTRTCT	NPAPL	NGGAF	CEGQSV	-QKIAC	TTLCPVDGRWTPW 373				
QY	306	SKWSAC	GLDCTH	MRSRCS	DPAPR	NGGEE	COGTD	LTRNCT	SDLCVHSAS	GEDVALYVG 365				
Db	374	SKWSTG	CTECHT	WRRECT	PAPKNG	GKDCD	GLVLQ	SKNCTD	GLCMQTAP	DDSDVALYVG 433				
QY	366	L-TAVAV	CLVLLV	LIVLIV	CRKKEG	LDSD	VADSS	ILTSQ	PGVPSIK	PKSKADNPHLLTIQ 424				
Db	434	IVIAVIV	CLAI	SVVVAL	IVYRKN	RHDFES	DIIDSS	ALNGG	PGVPSIKAA	RQD---LLAVP 490				
QY	425	PDL	-STTTT	VGSLCP	QDGP	SPKFQ	LTING	HLSP	LGGR	THLHS-----SPTSEAE	FEV 479			
Db	491	PDLT	SAAMV	KGPVYAL	HD-VSD	KIPMTNS	PILDP	PNLKI	KVNTSG	ANS	PQDDLEFT 549			
QY	480	SRLS	---TONYF	-----	RSLPRGT	--SNMTY	GTGFL	NGRIMI	PTNGIS	LLIPDDA 525				
Db	550	SKLSPQ	MTQSLE	NEALS	SLKNQ	SLARQ	TDPSCT	AFGFS	NSLGGH	LIVNSGV	SLIPAGA 609			
QY	526	IPROK	IEYI	TLHK	PDVRL	PLAGC	QTLLS	PIVSC	GPVGL	TRPVIL	AMDHG	GEPSD 585		
Db	610	IPOGRV	EMYV	YVHR	KETMR	PPMD	DSQ	TLLTF	VVSC	PPGALL	TRPVVIL	MHCADPNT 669		
QY	586	SWSLRL	KQSC	CEGSE	WEDV	HLGEP	SHLYY	COLEA	SACV	VFTEQ	LGREFAL	VEALSVA 645		
Db	670	DWKILL	KNQAQ	QGW	-D	VVVV	GENFT	PCYIK	LDAE	ACHILL	TENLSTYALV	GHSTKA 728		
QY	646	AAKRLK	LLFAP	ACTS	LEYNIR	YVCL	HDTH	DALK	EVVQ	LKQLG	QGLQ	LIQPRVILHFKDS 705		
Db	729	AAKRLK	LAIF	GPLCC	SSLEYS	IRVCL	DDTD	QDAL	KEIL	HLERQ	TGGQLLE	EPKALHFKGS 788		
QY	706	YHNLRL	SIHDV	PSLW	SKLLVS	QEI	PPYH	WNGT	QRYL	HTCTFL	ERVSP	STSDLACKL 765		
Db	789	THNLRL	SIHDV	PSLW	SKLLVS	QEI	PPYH	WNGT	QRYL	HTCTFL	ERVSP	STSDLACKL 848		
QY	766	WWOVED	GQGS	FNFT	TKDTR	AEAL	LALE	SEAG	VPAL	GVPSA	PKIPFL	RQKIISLD 825		
Db	849	CVROVE	GEGQ	FQ	LNCTV	SEFT	GIDL	PLDP	ANTIT	VTGSA	PSIDLP	ROKLCSSLD 908		
QY	826	PPCRG	ADWRT	LAQ	LHDS	LSLFS	FAKSP	SP	TAMIL	NLWEA	RHPNG	NLSQIAA	AVAGL 885	
Db	909	APQTRG	HDWR	MLA	HLK	NLD	RYNL	YVAT	KS	PTGV	ILDWEA	QNF	PDGNSL	MLAAVLEBMG 968
QY	886	QPDAG	LFTV	SEAE	898									
Db	969	RHETV	VS	LA	EBQ	981								

RESULT 13  
 ADE63096  
 ID ADE63096 standard; protein; 945 AA.  
 XX  
 AC ADE63096;  
 XX

DT	29-JAN-2004	(first entry)
XX		
XX	Rat Protein AAB57679, SEQ ID NO 9031.	
DE		
XX		
KW	Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;	
XX	chronic constriction injury; CCI; spared nerve injury; SNI; Chung.	
KW		
XX		
OS	Rattus norvegicus.	
XX		
WO	2003016475-A2.	
PN		
XX		
PD	27-FEB-2003.	
XX		
XX	14-AUG-2002; 2002WO-US025765.	
PF		
XX		
14	AUG-2001; 2001US-0312147P.	
PR		
PR	01-NOV-2001; 2001US-0346382P.	
XX		
PR	26-NOV-2001; 2001US-033347P.	
XX		
XX	(GEHO ) GEN HOSPITAL CORP.	
PA	(FARB ) BAYER AG.	
XX		
XX	Woolf C, D'urso D, Befort K, Costigan M;	
PI		
XX		
XX	WPI; 2003-268312/26.	
DR		
XX	GENBANK; AAB57679.	
XX		
PT	New composition comprising two or more isolated polypeptides, useful for	
PT	preparing a medicament for treating pain in an animal.	
XX		
XX	Claim 1; Page: 1017pp; English.	
PS		
XX		
XX	The invention discloses a composition comprising two or more isolated rat	
CC	or human polynucleotides or a polynucleotide which represents a fragment,	
CC	derivative or allelic variation of the nucleic acid sequence. Also	
CC	claimed are a vector comprising the novel polynucleotide, a host cell	
CC	comprising the vector, a method for identifying a nucleotide sequence	
CC	which is differentially regulated in an animal subjected to pain and a	
CC	kit to perform the method, an array, a method for identifying an agent	
CC	that increases or decreases the expression of the polynucleotide sequence	
CC	that is differentially expressed in neuronal tissue of a first animal	
CC	subjected to pain, a method for identifying a compound which regulates	
CC	the expression of a polynucleotide sequence which is differentially	
CC	expressed in an animal subjected to pain, a method for identifying a	
CC	compound that regulates the activity of one or more of the	
CC	polynucleotides, a method for producing a pharmaceutical composition, a	
CC	method for identifying a compound or small molecule that regulates the	
CC	activity in an animal of one or more of the polypeptides given in the	
CC	specification, a method for identifying a compound useful in treating	
CC	pain and a pharmaceutical composition comprising the one or more	
CC	polypeptides or their antibodies. The polynucleotide or the compound that	
CC	modulates its activity is useful for preparing a medicament for treating	
CC	pain (e.g. spinal segmental nerve injury (Chung), chronic constriction	
CC	injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene	
CC	therapy). The sequence presented is a rat protein (shown in Table 2 of	
CC	the specification) which is differentially expressed during pain. Note:	
CC	The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic form directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences.	
XX		
XX	Sequence 945 AA;	
XX		
XX		

Query Match	52.2%;	Score	2497;	DB	7;	Length	945;															
Best Local Similarity	52.0%;	Pred. No.	2.6e-207;																			
Matches	501;	Conservative	144;	Mismatches	236;	Indels	82;															
Gaps	20;																					
Qy	1	MAVRPGLWPALLGIVLA	AW	-----	LRG	--SGAQS	SATVANPVP	PGANPD	LLPHFL	VEP	BDV	53										
Db	1	MRASGARGALLLALL	LLCWDPT	SLAGID	SGQ	---	ALPDS	FP	GA	PA	QALPHFL	LEP	DA	57								
Qy	54	YIVK	NPVLLV	CKAV	PATQ	IF	FKNC	GEW	RVQD	VHIER	STDS	SGS	CEPT	MEV	IRV	NSROOV	113					
Db	58	YIVK	NPV	ELH	CR	AP	ATQ	IY	FKNC	GEW	VSQ	KGH	V	Q	ESL	DEAT	GLRIR	EV	Q	IRV	NSROOV	117



QY	1	MAVRPGLWPAALLGIVLAAM---	LRGSAQOASATV-ANPFCANPDLPHFLVPEPDIYV	56
Db	1	MGARSGARGALLALLCWDPRLSOAGRKSGEVLDPSPSAPAEPLPYFLQBPQDAYIV	60	
QY	57	KNKPVLLVCKAVPATQIFKCKNGEHWVQDVHVIETSDGSS-----CEPTMEVRINVSQ	111	
Db	61	KNKPEVLRCAFFATQIFYKCKNGEWSQNDHVTQEGLEDEATLGARGGLRVREVOIEVSRQ	120	
QY	112	QVEKVFGLGEEVWCOCVWSSGTTKSKAYTRIALRKNPEOEPLAKEVSLQEOQVILPCR	171	
Db	121	QVEELFLEEDYWCOCVWSSAGTTKSRAYVRIACLRKNFDOEPLKEVPLDHEVLQCR	180	
QY	172	PPGIPPAEVEWLRNEDVLPSPDPNVIYITREHSLVVRQARLADANTYCVAKNIVARR	231	
Db	181	PPGVPVAEVEWLNKNEVDIPTQDTNFTLLTIDHNLIRIARLSDTANTYCVAKNIVARR	240	
QY	232	SASNAVIYVNGWSWTWNSVCASCGRGWOKRSCTNPAPLNGGAFCEGQNVHDTV	291	
Db	241	STTATVIYVNGWSSWAENSPCNRGCGKRTCTNPAPLNGGAFCEGQ-AFOKTA	299	
QY	292	SSILVSDGSPWSPKWSACGLDCTHWRSCSDPAPRNGGECQGTDLTRNCTSDLCV	351	
Db	300	CTTICPDGATWENKWSACSTCAHWRSCRECMAPPQNGGRCDSGLTLLSKNCTDGLCM	359	
QY	352	HS-----ASGPEDVALYVGL-IAVAVCIWLLVLLVILVYCKKGLSDVADSS-IL	401	
Db	360	QSEFPVPAVLEASG--DAALYAGLVAVFVVAAILMAVGVVYRNRCDFTDITDSSAAL	417	
QY	402	TSGPQPSIKSKADNPILL--TTQPDLSITT--TYQSLCPROGSPKFOITNGHLISL	458	
Db	418	TGGFHPVNFKTPSPNPQLLPSPPDLTASAGYRGVYALQDS-TDKIPMTNSPLLD	476	
QY	459	LGGRHRLTHHSPT-----SEAEFVSRLSTQNY-----FRS-----	490	
Db	477	LPSLKVYVSSSTGSGFLADGADLLGLVPPGYPSDFARDTHFLHRSASLSGSQLLG	536	
QY	491	LPRGTSNMTYGFNPLGRLMIPNTGISLLIPDPAIPRGKIYIEYILTHKPEDVRLPLA-	549	
Db	537	LPRDPGSSVSGTFGCLGRLSIPGTGVSLLVPGAIPQCKFYEMVYLLINKAEST-LPLSE	595	
QY	550	GCOTLLSPIVSCGPPGVLLTPVILAMDHCGEPSPDSWLSRLKKSCSGSEQDVLHGE	609	
Db	596	GTQTVLSPSVTCGPTGLLLCRPVILTMPHCAEVSARDWIFQKTOAHQHQWEQEVVTLDE	655	
QY	610	EAPSHLYCOLEASACVYFTQLGRFALVGEALSVAARLKLKLLFAPVACTSLEYNIRV	669	
Db	656	ETLNTPCYQLEPRACHILLDQLTGYVFTGESYSRAVKRLQAVFAPALCTSLEYSRV	715	
QY	670	YCLDHTDALKEVVOLEKQGLQIQRVRLHFKDSYNLRLSHIDVPSSLWKSLLVSY	729	
Db	716	YCLEDTVPALKEVLELETLGGLVVEEPKLMFKDSYNLRLSHLDLPHAWRSKLLAKY	775	
QY	730	QEIPEYHLWGTORVILHCTFLERSVPSSTDLACKLWQVGEQGSIFNFINIKTRFF	789	
Db	776	QEIPEYHLWGSQKALHCTFLERSHLSASTELTCKICVRQVEGEQIIFQLHTTLA-ETPA	834	
QY	790	AELIALESEAG--VPALVGPFAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLDHSL	847	
Db	835	GSLLDLCAPSGSTVTLQGPYAFKIPLSIRQICNSLDAPNSRGNDRMRLAQLKLSMDRYL	894	
QY	848	SFFASKPSPKTAMILNWEARHPNGLSQLAAVAGLQOPDAGLFTVSEAE	899	
Db	895	NYFATKASPTGVILDLEALQDDGLNSLASALEMGKSEMLVAVATDGD	946	

RESULT 15  
AAW78900  
ID AAW78900 standard; protein; 943 AA.  
XX  
AC AAW78900;

XX	25-MAR-2003	(revised)	
DT	21-DEC-1998	(first entry)	
XX	Rat UNC-5 homologue UNC5H-2.		
DE	UNC5H-2; rat; netrin receptor; cell migration; axon guidance;		
DE	diagnosis; therapy.		
XX	Rattus sp.		
XX	Key	Location/Qualifiers	
XX	Peptide	148..161	
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
FT	Misc-difference	753	
FT	/note= "encoded by CG"		
FT	Peptide	909..924	
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
XX	W09837085-A1.		
XX	27-AUG-1998.		
XX	19-FEB-1998;	98WO-US003143.	
XX	19-FEB-1997;	97US-00808982.	
XX	(REGC ) UNIV CALIFORNIA.		
XX	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Kleinomasu K;		
XX	WPI: 1998-495364/42.		
XX	N-PSDB; AAV52942.		
XX	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and		
XX	the biopharmaceutical industry.		
XX	Claim 1; Page 24-26; 32pp; English.		
XX	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis		
XX	elegans UNC-5 protein. Their amino acid sequences were deduced from		
XX	isolated unc5H cDNA clones (see AAV52940 and AAV52942) isolated from an		
XX	E18 brain cDNA library. The predicted proteins show similarity with UNC-		
XX	5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin		
XX	type-1 repeats, a predicted membrane spanning region, and a large		
XX	intracellular domain. They are predicted to be involved in cell migration		
XX	and axon guidance, and are characterised as receptor proteins for		
XX	netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins		
XX	are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly		
XX	from transfected host cells. The invention also provides unc-5		
XX	hybridisation probes and primers, vertebrate UNC-5-specific binding		
XX	agents such as specific antibodies, and methods of making and using the		
XX	subject compositions in diagnosis (e.g. genetic hybridisation screens for		
XX	vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate		
XX	vertebrate unc-5 gene expression) and in the biopharmaceutical industry		
XX	(e.g. as immunogens, reagents for modulating cell guidance, reagents for		
XX	screening chemical libraries for lead pharmacological agents, etc.).		
XX	(Updated on 25-MAR-2003 to correct PI field.)		
XX	Sequence 943 AA;		
XX	Query Match	52.0%; Score 2490; DB 2; Length 943;	
XX	Best Local Similarity	52.3%; Pred. No. 1.1e-206;	
XX	Matches	496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;	

QY	9	PALGIVLAAMLGSCAQOASATVNPVPGANPDLPHFLVPEPDIYVKNKPVLLVCKAV	68
Db	21	PSLAGI-----DSGAQ---GLPDSFPFAEQPLFLEPEDAYIVKNKPVLLHCRAF	70
QY	69	PATQIFFKCKNGEHWVQDVHVIETSDGSGGETMEVRINVSQQVEKVFGLBEYWCQVVA	128
Db	71	PATQIFYKCKNGEWSQKHVTKQESLDEATGLRIVQIEVSRQVVEELFGLDYWCQVVA	130





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 6, 2004, 18:20:44 ; Search time 152 Seconds  
(without alignments)  
3282.242 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPLGLWALLGIVLAAML.....AVAGLQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh

-Q=/cgn2\_1/USPTO.spool/US09970944/runat\_05102004\_112011\_1655/app\_query.fasta\_1.1095

-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09970944 @CGN\_1\_1\_75 @runat\_05102004\_112011\_1655 -NCPUS=6 -ICPU=3

-NO\_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:\*

1: /cgn2\_6/prodata/2/ina/5A COMB.seq:\*

2: /cgn2\_6/prodata/2/ina/5B COMB.seq:\*

3: /cgn2\_6/prodata/2/ina/6A COMB.seq:\*

4: /cgn2\_6/prodata/2/ina/6B COMB.seq:\*

5: /cgn2\_6/prodata/2/ina/PCTUS COMB.seq:\*

6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4545.5	95.0	3014	2	US-08-808-982-1
2	4545.5	95.0	3014	3	US-09-306-902A-1
3	2694	56.3	1787	2	US-08-808-982-2
4	2694	56.3	1787	3	US-09-306-902A-2
5	2463	51.5	2831	2	US-08-808-982-3
6	2463	51.5	2831	3	US-09-306-902A-3
7	2039	42.6	1282	4	US-09-833-381-1806
8	1123	23.5	1605	4	US-09-833-381-1807
9	600.5	12.5	771	1	US-08-253-155A-17
10	267.5	5.6	5749	4	US-09-262-537-48
11	253	5.3	305	2	US-08-808-982-4
12	253	5.3	305	3	US-09-306-902A-4

13	250	5.2	1326	3	US-08-985-526-4	Sequence 4, Appli
14	250	5.2	7231	4	US-09-919-172-64	Sequence 64, Appli
15	245.5	5.1	657	3	US-08-985-526-2	Sequence 2, Appli
16	211.5	4.4	3874	4	US-09-877-730-31	Sequence 31, Appli
17	204.5	4.3	3991	4	US-08-506-296B-3	Sequence 3, Appli
18	204	4.3	3210	4	US-09-877-730-1	Sequence 1, Appli
19	204	4.3	3453	4	US-09-877-730-7	Sequence 7, Appli
20	192.5	4.0	2715	4	US-09-877-730-5	Sequence 5, Appli
21	192.5	4.0	2958	4	US-09-877-730-9	Sequence 9, Appli
22	186	3.9	1143	4	US-09-877-730-3	Sequence 3, Appli
23	184	3.8	7702	4	US-09-023-655-1336	Sequence 1336, Ap
24	178.5	3.7	3675	4	US-09-930-872-3	Sequence 3, Appli
25	178.5	3.7	4042	4	US-09-930-872-5	Sequence 5, Appli
26	175.5	3.7	8312	4	US-09-620-312D-1048	Sequence 1048, Ap
27	173.5	3.6	6814	4	US-09-484-970B-66	Sequence 66, Appli
28	170.5	3.6	3126	4	US-09-392-184-7	Sequence 7, Appli
29	170.5	3.6	5690	2	US-08-447-464-2	Sequence 2, Appli
30	170.5	3.6	5690	2	US-08-716-679-2	Sequence 2, Appli
31	168	3.5	2114	4	US-09-130-491-7	Sequence 7, Appli
32	168	3.5	4548	4	US-09-571-479C-5	Sequence 5, Appli
33	167.5	3.5	3638	4	US-09-369-364A-8	Sequence 8, Appli
34	166	3.5	4192	4	US-09-122-126B-1	Sequence 1, Appli
35	166	3.5	4192	4	US-09-634-286A-1	Sequence 1, Appli
36	165	3.4	4843	3	US-08-986-485-1	Sequence 1, Appli
37	164.5	3.4	3218	4	US-09-369-364A-6	Sequence 6, Appli
38	163	3.4	4078	4	US-09-016-434-1132	Sequence 1132, Ap
39	163	3.4	6000	1	US-08-348-006B-6	Sequence 6, Appli
40	163	3.4	6000	2	US-08-800-825A-6	Sequence 6, Appli
41	163	3.4	6000	3	US-09-158-657-6	Sequence 6, Appli
42	163	3.4	6000	5	PCT-US94-10166-6	Sequence 6, Appli
43	162.5	3.4	5357	4	US-09-392-184-5	Sequence 5, Appli
44	160.5	3.4	3889	4	US-09-568-559-1	Sequence 1, Appli
45	160.5	3.4	4676	4	US-09-130-491-1	Sequence 1, Appli

#### ALIGNMENTS

##### RESULT 1

US-08-808-982-1

; Sequence 1, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/808,982

; FILING DATE:

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC96-217

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3014 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-808-982-1

Alignment Scores:

Pred. No.: 0 Length: 3014  
 Score: 4545.50 Matches: 852  
 Percent Similarity: 96.78% Conservative: 19  
 Best Local Similarity: 94.67% Mismatches: 26  
 Query Match: 94.66% Indels: 3  
 DB: 2 Gaps: 3

US-09-970-944-2 (1-899) x US-08-808-982-1 (1-3014)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeuValLeuAlaAlaTrpLeu 20  
 DB 1 ATGGCGGTCCGGCCCGCCCTGTGGCCAGTGTCTCTGGGCATAGTCTCTGGCCGCTGGCTT 60  
 QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
 DB 61 CGTGGTTCGGGTGCCCGAGAGTGCACAGGTGGCCCAATCCAGTGCCTGGTGCACACCCC 120  
 QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60  
 DB 121 GACCTGTGCCCTTCTCTGGTAGAGCTGTAGAGCGTGTACATTTGTCAAGAACAAAGCCG 180  
 QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
 DB 181 GTGTGTGTGTGTGCAAGGCTGTGCTCCCTGCCACCCAGATCTCTTCAAGTGCATGGGAA 240  
 QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
 DB 241 TGGTCCGCCAGTTCGATCAGTAATTTGAACGACAGCACCGACAGCAGCGGATTGCCA 300  
 QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGlnLysValPheGlyLeuGlu 120  
 DB 301 ACCATGAGGTCCGTATCAACGTATCGAGGCAGCAGTAGAGAAAGTGTGGGCTGGAG 360  
 QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140  
 DB 361 GAAATCTGTGTGCCAGTGTGTGGATGAGTCTCTCGGTACACCAAAAGTCAAGAGCC 420  
 QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
 DB 421 TACATCCGATTGCTTATTTGGCAAGAACTTTTGGACGAGGAGCCACTGGCCAAAGAGTG 480  
 QY 161 SerLeuGlnGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180  
 DB 481 TCATGAGCAAGCAATGTACTACCTTGTGCCCCCGCAAGAAATCCCGCCAGCTGAG 540  
 QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
 DB 541 GTGAGTGGCTTCGAAATGAGGACCTCTGTGACCCCTCTCTCGATCCCAATGTGTATC 600  
 QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
 DB 601 ACGCGGAGCACAGCTTAGTGTGTGCGTCAAGCGCGCTGGCCGACACGCGCAATACACC 660  
 QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
 DB 661 TGTGTGGCCCAAGAACATCGTAGCCCGTGGCAAGACCTCTGAGCGGTCATGTGTAT 720  
 QY 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
 DB 721 GTNACCGGTGGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780  
 QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyAlaPhe 280  
 DB 781 GGCTGGCAGAAACGAGCGCGAGTGCACCAACCGGCACCTCTCAACGCGGGCGCTTC 840

QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
 DB 841 TGTGAGGGGAGATGTC---CAGAAACAGCGCTGGCCACTCTGTGCCAGTGGATGG 897  
 QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
 DB 898 AGCTGGAGTTCGTGGAGTAAGTGTGAGCTGTGGGCTTGACTGCACCCACTGGCGGAGC 957  
 QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
 DB 958 CGCGAGTGTCTGACCCAGCACCCGCAATGGAGGTGAGAGTGTGCGGGTGTGACCTG 1017  
 QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
 DB 1018 GACCCCGCACTGTACCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
 QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
 DB 1078 GCTCTCTACATGGCCCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1137  
 QY 381 IleLeuValTyrCysArgLysLysGlyGlyLeuAspSerAspValAlaAlaAspSerIle 400  
 DB 1138 GGACTCATTTACTGTGCGAAGAGGAGGCTGACTCCGATGTGCGGACTGCTGCATC 1197  
 QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAlaAspAsnProHisLeu 420  
 DB 1198 CTCACCTCGGGCTTCAGCTGTGACATCAAGCCAGCAAGACAGACACCCACCTG 1257  
 QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlySerLeuCysPro 439  
 DB 1258 CTCACCATCCAGCCAGCTCAGCACCACTTACCATCTACCGGCGAGTCTATGTTCG 1317  
 QY 440 ArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisIleLeuSerProLeu 459  
 DB 1318 AGCAGAGATGGACCCAGCCCAAGTTCAGCTCTAATGGTCACTGTCTGAGCCACTG 1377  
 QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479  
 DB 1378 GGGAGTGGCGGCATAGTTGCACAGCTCACCCACCTCTGAGGCTGAGGACTTGTGTC 1437  
 QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
 DB 1438 TCCCGCTCTCCACCCCAAACTACTTTCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1497  
 QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
 DB 1498 TACGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACGGGATCAGCCTC 1557  
 QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539  
 DB 1558 CTATACCCCGGATGCCATCCCGGAGGAGGATCTACGAGATCTACCTCACACTGCAC 1617  
 QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
 DB 1618 AAGCCAGACACGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1677  
 QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
 DB 1678 AGCTGTGGGCGCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1737  
 QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599  
 DB 1738 GGAGAGCCAGCCCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1797  
 QY 600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619  
 DB 1798 TGGGAG---CATGTGTGTGACCTTGGTGGAGGAGTACCTTCCCACTCTACTACTGCCAG 1854  
 QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
 DB 1855 CTGAGGCGCGGGCGCTGTATGTCTTCCAGAGCAGCTGGCGCGCTTTCCTGCTGCTGCTG 1914

```
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCTCTTGTTGCTCCGGTGGCC 1974
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTAAGTCCCTTGAGTACAACTCCAGTGTACTGCTTACAGACACCCACGAGCTCTC 2034
QY 680 LysGluValValGlnLeuLysGlnLeuGlyGlnLeuIleGlnLeuProArgVal 699
Db 2035 AAGAGGTGGTGCAGCTGAGAGAGAGCTAGGTGGACAGCTATCCAGAGGCTCGCGTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACCTCAAAGACAGATTACCAACCTACGCTCTCCATCCAGCGTGGCCAGCTCC 2154
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAAGAGCAAGCTACTGTCTGACTACAGGAGATCCCTTTTACCACATCTGGAAC 2214
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGCATCTGCACTTGCACTTCACTTCCCTGGAGCGCATCAACGCCAGCACCAGC 2274
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGGCTGCAAGGTGGTGGTGGCAGGTGGAGGAGATGGGAGAGCTTCAACATC 2334
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2335 AACCTTCAACATCACTAAGGACACAAGGTTTGTGAATTGTTGGCTCTGGAGAGTGAAGG 2394
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCACAGCCCTGGTGGCCCGAGTGCCTTCAAGATCCCTTCTATTCCGCCAAGAG 2454
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATCATCGCCAGTCTGGACCCACCTTCAGCCGGGGCGGAGTGAAGAACTTAGCCGAG 2514
QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2515 AAACCTTCACTGGACAGCATCTTAGCTTCTTTGGCTTCAAGCCAGCCCTACAGCCATG 2574
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaIa 879
Db 2575 ATCTTCACTATGGAGGACGAGCATTTCCCAACGGCAACTTCGGCAGGCTGGCAGCA 2634
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2635 GCTGTGGCCGAGTGGGCCAACACAGATGCTGGCCTTCTCACGGTCTCGGAGGCCAGTGT 2694

RESULT 2
US-09-306-902A-1
; Sequence 1, Application US/09306902A
; Patent No. 6277585
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/306,902A
; FILING DATE: 07-May-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-306-902A-1

Alignment Scores:
Pred. No.: 0 Length: 3014
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: 3 Gaps: 3

US-09-970-944-2 (1-899) x US-09-306-902A-1 (1-3014)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db 1 ATGGCCGTCCGGCCCGGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTGCCGCTTGGCTT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 61 CGTGTTCGGGTGCGCCAGCAGAGTCCACGGTGGCCAATCCAGTCCCGGTGCCAACCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db 121 GACCTGTGCCCCACTTCTGTAGAGCTGAGGACGTGATCTTCAAGTGCATTCGGGNA 180
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 181 GTGTGTTGGTGTCAAGGCTGTGCTGCCACCCAGATCTTCTCAAGTGCATTCGGGNA 240
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db 241 TGGGTCCGCCAGTGCATCAGCTAATTGAACGCAGCACCGACAGCAGCGGATTGCCA 300
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 301 ACCATGGAGGTCGGTATCAAGTATCGAGGCGAGGAGTAGAGAAAGTGTTCGGCTGGAG 360
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
Db 361 GAATACTGGTCCAGTGTGTGGCATGGAGCTCTCTCGGATACCAACCAAAAGTCAGNAGGCC 420
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 421 TACATCCGATTGGCTATTTCGCAAGAACTTTGAGCAGGAGGCACCTGGCCCAAGAAAGTG 480
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
Db 481 TCACTGGAGCAAGGCATTTGTACTACCTTGTGCCGCCCCCGAGAAAGAAATCCCCCGCTGAG 540
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
Db 541 GTGGAGTGGCTTCGAAATGAGACCTGTGGACCCCTCCCTCGATCCCATGTGTATCATC
```



APPLICANT: Tessier-Lavigne, Marc  
 APPLICANT: Leonardo, E. David  
 APPLICANT: Hink, Lindsay  
 APPLICANT: Masu, Masayuki  
 APPLICANT: Kazuko, Keino-Masu  
 TITLE OF INVENTION: Netrin Receptors  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 BUSH STREET, SUITE 3200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/808,982  
 FILING DATE:  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OSMAN, RICHARD A  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: UC96-217  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 343-4341  
 TELEFAX: (415) 343-4342  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1787 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-08-808-982-2

## Alignment Scores:

Pred. No.: 8,576-252 Length: 1787  
 Score: 2694.00 Matches: 543  
 Percent Similarity: 97.14% Conservative: 1  
 Best Local Similarity: 96.96% Mismatches: 11  
 Query Match: 56.28% Indels: 11  
 DB: 2 Gaps: 2

US-09-970-944-2 (1-899) x US-08-808-982-2 (1-1787)

344 AsnCysThrSerAspLeuValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363  
 3 AACTGTACAGTACCTCTG-GTACACACTGCTTCTGGCCCTGAGGACGTGGCCCTCTAT 61  
 364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuVal 383  
 62 GTGGGCTCATCGCGTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121  
 384 TyrCysArgLysLysGlyLeuAspSerValAlaAspSerSerIleLeuThrSer 403  
 122 TATTGCCGAAGAAGAGGGGCTGACTCAGATGGCTGACTCGTCCATTCTACCTCA 181  
 404 GlyPheGlnProValSerIle-LysProSerIysAlaAspAsnProHisLeuThrIle 423  
 182 GCTTCCAGCCCGTCTGAGCATTAAGCCAGCAAGACACACCCCTGCTGCTGCTGCT 241  
 423 eGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThrThrThr 442  
 242 CCAGCCGACCTCAGCACCAACACACCTACAGAGGAGCTCTCTGCTCCCGGAGGA 301  
 442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 462  
 302 TGGGCCAGGCCCAAGTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGG 361

462 YArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgIle 482  
 362 CGCCACACACTGCACACACAGCTCTCCACCTCTCTAGGCCGAGGAGTTCGTCTCCGCT 421  
 482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetTyrTyrGlyTh 502  
 422 CTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGACACCAACATGACCTATGGGAC 481  
 502 rPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuPr 522  
 482 CTTCAACTTCTCTCGGGGCGCGCTGATGATCCTAATAACAGGAATCAGCCTCTCATCCC 541  
 522 cProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGl 542  
 542 CCCAGATGCCATACCCCGAGGAGAGATCTATGAGATCTACCTCAGCTGCACCAACCCGGA 601  
 542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGl 562  
 602 AGAGTGAGGTTCGCCCTAGCTGGCTGTTCAGACCTGCTGAGTCCCATCGTTAGCTGG 661  
 562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582  
 662 ACCCTCGCGCTCTGCTCACCCGCGCAGTATCTGCTATGGACCACTGTGGGGAGGC 721  
 582 oSerProAspSerTyrTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTyrGluGl 602  
 722 CAGCCCTGACAGCTGGAGCTGGCCCTCAAAAGCAGTCGTGCGAGGAGGC-TGGAG-- 778  
 602 nAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622  
 779 -GATGT-CTGCACCTGGCGGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAGGC 836  
 622 aserAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaIle 642  
 837 CAGTGCCTGTACGTCTTTCACCGAGCAGCTGGCGCGCTTTCCTGCTGGGAGAGCCCT 896  
 642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSe 662  
 897 CAGGTGGCTCGCCGCAAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCGCTGCACCTC 956  
 662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682  
 957 CCTCGAGTACACATCCGGGTCTACTGCTGATGACACCCACCATGATGACCTCAAGAGAGT 1016  
 682 lValGlnLeuGlyLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPh 702  
 1017 GGTGCAGCTGGAGAAGCAGCTGGGGGAGCAGCTGATCCAGGAGCCACGGGTCTGCACTT 1076  
 702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTyrPly 722  
 1077 -AAGGACAGTTACCACACCTGCC-CTATC-ATCCAGATGTGCCAGCTCCCTGTGGAA 1133  
 722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTyrAsnGlyThrGl 742  
 1134 GAGTAAGCTCTTGTGCTGACCTACAGGAGATCCCTCTTTATCACATCTGGAATGGCACGCA 1193  
 742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762  
 1194 CGCGTACTTGCATCGACCTTCACCTGGAGCGTGTGAGCCCGGAGCAGCTAGTGACCTGGC 1253  
 762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782  
 1254 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGAGCGGAGCAGAGCTTCAGCATCAACTTCAA 1313  
 782 nileThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802  
 1314 CATCACCAAGACACACAGGTTTGTCTGAGCTGCTGGCTCTGGAGAGTGAAGGGGGGTCCC 1373  
 802 oAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSe 822  
 1374 AGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTATTCCGAGAGATTAATTC 1433  
 822 rSerLeuAspProProCysArgArgGlyAlaAspTyrArgThrLeuAlaGlnLysLeuHi 842





```
QY 722 sSerLysLeuLeuValSerTyrGlnGlnLeuProPheTyrHisIleTyrAsnGlyThrGl 742
Db 1134 GAGTAAGCTCTGTGTAGCTACAGGAGATCCCTTTTATCATCATCTGGAATGCACGCA 1193
QY 742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
Db 1194 GCGGTACTTGGCACTGACCTTACCTGAGCGTGTGAGCGGTGTGAGCGGAGCTAGTACCTGGC 1253
QY 762 aCysLysLeuTyrValTyrGlnValGlnGlyAspGlyGlnSerPheSerIleAsnPheAs 782
Db 1254 CTGCAAGCTGTGGTGTGGAGGTGAGGGCGGAGCGGAGCTTCAGCATCAACTTCAA 1313
QY 782 nIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValPr 802
Db 1314 CATCAACAGGACACAAAGGTTTGCTCAGCTGTGGTCTGTGGAGGTGAAGCGGGGTCCC 1373
QY 802 oAlaLeuValGlyProSerAlaPheIleLysIleProPheLeuIleArgGlnLysIleIleSe 822
Db 1374 AGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGGAATATTTTC 1433
QY 822 rSerLeuAspProProCysArgArgGlyAlaAspTyrArgThrLeuAlaGlnLysLeuHi 842
Db 1434 CAGCTGGACCAACCTGTAGCGGGTCCGACTGGCGGACTCTGGCCAGAACTCCA 1493
QY 842 sLeuAspSerHisLeuSerPhePheAlaSerLysLysProSerProThrAlaMetIleLeuAs 862
Db 1494 CTTGGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCGCCACAGCCATGATCTCAA 1553
QY 862 nLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAl 882
Db 1554 CTTGTGGGAGGCGGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGC 1613
QY 882 aGlyLeuGlyGlnProAspAlaGlyLeu-PheThrVal-SerGluAlaGluCys 899
Db 1614 TGGGACTGCCAGCAGGACGGTGGCTTTTTCACAGTGTTCGAGGCTGAGTGC 1667

RESULT 5
US-08-808-982-3
; Sequence 3, Application US/08808982
; Patent No. 5939271
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Leonardo, E. David
; APPLICANT: Hink, Lindsay
; APPLICANT: Masu, Masayuki
; APPLICANT: Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 2831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-808-982-3
```

## Alignment Scores:

```
Pred. No.: 5,72e-229 Length: 2831
Score: 2463.00 Matches: 495
Percent Similarity: 67.41% Conservative: 144
Best Local Similarity: 52.22% Mismatches: 227
Query Match: 51.45% Indels: 83
DB: 2 Gaps: 19
```

US-09-970-944-2 (1-899) x US-08-808-982-3 (1-2831)

```
QY 9 ProAlaLeuLeuGlyIleValLeuAlaAlaTyrLeuArgGlySerGlyAlaGlnGlnSer 28
Db 61 CCGAGCTTAGCAGGCATT-----GACTCTGGTCCCGAG----- 93
QY 29 AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHisPheLeuVal 48
Db 94 ---GGACTCTCCAGACTCCTTCCATCAGCACCCGCGGAGCAGCTGCTCACTTCCTGCTG 150
QY 49 GluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaVal 68
Db 151 GAACCCAGAGGATGCGCTACATCGTAAAGAACAAAGCCAGTGAATTCACCTGCCGAGCCTTC 210
QY 69 ProAlaThrGlnIlePhePheLysCysAsnGlyGluTyrValArgGlnValAspHisVal 88
Db 211 CTGCGCACACAGATCTACTTCAAGTGTAAATGGCGAGTGGTTAGCCAGAAAGCCACGCTC 270
QY 89 IleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnVal 108
Db 271 ACGCAGGAGAGCTCGATGAGGCCACAGCTTGGCGAATACGAGAGGTGCAGATAGAGGTG 330
QY 109 SerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTyrCysGlnCysValAla 128
Db 331 TCGCGCAGCAGCGGTGGAGGAACCTTTTGGGCTCGAGGACTACTGGTGTCACTGGCTGGCC 390
QY 129 TrpSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArg 148
Db 391 TGGAGCTCTTCGGGNAACACCAAGAGTCGCCAGCTACATCCGCATTGCCCTACTTGGCC 450
QY 149 LysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeu 168
Db 451 AAGAACTTTGACCAGAGCCCTCTGGCGAAGGAGGTACCCCTTGGATCATGAGGTCTCTCTG 510
QY 169 ProCysArgProProGluGlyIleProProAlaGluValGluTyrLeuArgAsnGluAsp 188
Db 511 CAGTGCCGCCCCACAGAGGGAGTGCCTGTGCTGAGGTGGAATGGCTCAAGAATGAAGT 570
QY 189 LeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValVal 208
Db 571 GTCATCGATCCGCTCAGGACACTTAACCTCTCTGCTCACCATTGACCACCACTCATCATC 630
QY 209 ArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAla 228
Db 631 CGCCAGGCGCGCTCTCAGACACAGCCCACTACACCTGTGTGGCAAGAATAATTGTGGCC 690
QY 229 ArgArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyIleTyrSerThrTyr 248
Db 691 AAGCGCGGAGCAGCAGCGCCACAGTCTATGTGTGAACGAGGTTGGTCCAGCTGG 750
QY 249 ThrGluTyrSerValCysSerAlaSerCysGlyArgGlyTyrGlnLysArgSerArgSer 268
Db 751 GCAGATGGTCAACCTGCTTAACCGCTGGCGCGGAGTTGGCAGAAACGCTACTAGGACC 810
QY 269 CysThrAsnProAlaProLeuAsnGlyIleAlaPheCysGluGlyGlnAsnValHisAsp 288
Db 811 TGCACCAACCCAGCCCCACATCAATGGAGGTGCTCTTCGAGGAGGACAG---GCTTGGCCAG 867
```

QY 289 ArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerIysTrp 308  
D5 868 AAGACGGCTTGACACCGCTGCGCCAGTGTAGTGGAGCTGAGTGGAGCAAGTGG 927  
QY 309 SerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaPro 328  
D5 928 TCCGCTGAGCAGACAGAGTGTGCGACCTGGCGCAGCCGCGAGTGTATGGCACCGCGGCC 987  
QY 329 ArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAsp 348  
D5 988 CAGACGGAGGCGGTGACTGACGAGCGGAGCTTACTTGTCTCAAGAACTGACCGATGG 1047  
QY 349 LeuCysVal-----HisSerAlaSerGlyProGlu----- 358  
D5 1048 CTGTGCTGCTGAATCAGAGAACTCTAAACGACCCCTAAAGCGCGCCCTGGAGCGGTG 1107  
QY 359 ---AspValAlaLeuTyValGlyLeu---IleAlaValAlaValCysLeuValLeu 376  
D5 1108 GGAGACGTGGCGCTGTATGCGGGCTCTGGTGGCGCTCTTGTGGTCTGGAGTTCTC 1167  
QY 377 LeuLeuValLeuIleLeuValTyCysArgLysLysGluGlyLeuAspSerAspValAla 396  
D5 1168 ATGGCTGTAGGAGTGTATGACCGAGAACTGCGGGACTTGCACACGAGACACT 1227  
QY 397 AspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla 415  
D5 1228 GACTCTCTGCTGCTCCTCACTGTGTGTCTTCCACCGCTCAACTTCAAGACTGCAAGGCC 1287  
QY 416 AspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThrThr---Thr 432  
D5 1288 AGCAACCCACAGCTCTGACCCACTCCGCCCTCCGGACTTACGGCCAGTGTGGCATC 1347  
QY 433 TyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsn 452  
D5 1348 TACCGCGGACCTGTGTATGCTCCCTGCAGGACTCT---GCCGACAAAGATCCCTATGACTAAT 1404  
QY 453 GlyHisLeuLeuSerProLeu----- 459  
D5 1405 TCACCCCTTCTGGATCCCTTGCCAGCCTCAAGATCAAGTCTTATGACTCCAGCACCATC 1464  
QY 460 -----GlyGlyGly----- 462  
D5 1465 GGCTCTGGGGCTGGCTGGCTGTATGGAGCGGACTGCTGGGTCTTACACCGGTACA 1524  
QY 463 -----ArgHisThrLeuHisHisSerSerProThrSerGlu 474  
D5 1525 TACCCAGCGGATTTCTCCGGGACACCCACTTCTGCACCTGCGCAGC----- 1572  
QY 475 AlaGluGluPheValSerArgLeuSerThrGlnAsnTyPheArgSerLeuProArgGly 494  
D5 1573 -----GCCAGCCTTGTGTCCAG---CACCTCTGGGCTTCTCCAGAC 1614  
QY 495 ThrSerAsnMetThrTyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsn 514  
D5 1615 CCAGCAGAGAGTGTAGTGCACCTTGTGTGCTGGTGGAGGCTGACCATTCGCGC 1674  
QY 515 ThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyIleTyGluIle 534  
D5 1675 ACAGGGTCAAGCTGTGTGTACCAAAATGGAGCAATTCGCCAGGCGAAGTCTTATGACTG 1734  
QY 535 TyThrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThr 553  
D5 1735 TATCTAGTATCAACAGACTGAAGAGCACC---CTCCCACTTTCGGAGGTTCACAGACA 1791  
QY 554 LeuLeuSerProIleValSerCysGlyProGlyValLeuLeuThrArgProValIle 573  
D5 1792 GTATTGAGCCCTCGGTGACTGCTGGGGCCCAACGGGCTCTCTGTGCGCCCTGTGTGTC 1851  
QY 574 LeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLys 593  
D5 1852 CTCAGTGTGCCCACTGTGTGTAAGTCAATTCGCCGAGACTGGATCTTCCAGCTCAAGACC 1911  
QY 594 GlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluAlaProSer 613

D5 1912 CAGGCCCATCAGGGCCACTGGAG---CAGTGGTGACTTTGGATGAGGAGACTCTCGAAC 1968  
QY 614 HisLeuTyTrpCysGlnLeuGluAlaSerAlaCysTyTrpValPheThrGluGlnLeuGly 633  
D5 1969 ACCCCCTGCTACTGCTAGCTAAGCTAAATCCTGCCACATCCTGTTGGACCAGCTGGGT 2028  
QY 634 ArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeu 653  
D5 2029 ACCTACGTGTTTACGGGCGAGTCTTACTCCCGCTCCGAGTCAAGCGGTCCAGCTAGCC 2088  
QY 654 LeuPheAlaProValAlaCysThrSerLeuGluTyTrpAsnIleArgValTyTrpCysLeuHis 673  
D5 2089 ATCTTGGCCCGCCAGCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCTACTGTCTGGAG 2148  
QY 674 AspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeu 693  
D5 2149 GACACTCTGCGACACTGAAGGAGTCTTAGAGTGGAGGAGACTCTGGGTGGCTACTTGG 2208  
QY 694 IleGlnGluProArgValLeuHisPheLysAspSerTyTrpHisAsnLeuArgLeuSerIle 713  
D5 2209 GTGGAGGAGCCCAAGACTTGTCTTTAAGGACAGTTACCACTACCTAGCTC-TCCCTC 2267  
QY 714 HisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyTrpGlnGluLeuPro 733  
D5 2268 CATGACATCCCATGCGCCACTGGAGGAGCAACTACTGGCCCAAGTACCAGGAGATTCCC 2327  
QY 734 PheTyHisIleTrpAsnGlyThrGlnArgTyTrpLeuHisCysThrPheThrLeuGluArg 753  
D5 2328 TTCTACCATGTGGAAACGGCAGCAAGGCGCTGCACTGCACCTTTCACCTGGAGAGA 2387  
QY 754 ValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAsp 773  
D5 2388 CATAGCTAGCTTCCACTGAGTTCACCTGTAGTCTGCTGCGGAGGTAGAGGGGAA 2447  
QY 774 GlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeu 793  
D5 2448 GCGCAGATTTCCAGCTGCACACCAGCTGGCT---GAGACGCTCTGCTCCCTGGAT 2504  
QY 794 AlaLeuGluSerGluAlaGlyValProAlaLeu-----ValGlyProSerAlaPheLys 811  
D5 2505 GCACCTCTGCTCTGCCCTGGCAATGTGTCACACACACAGCTGGGACCTTATGCTTCAAG 2564  
QY 812 IleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGly 831  
D5 2565 ATACCACTGTCCATCCGCGAGAGATCTGCACAGCTTGACGCCGCCCACTCACGGGC 2624  
QY 832 AlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAla 851  
D5 2625 AATGACTGGCGGTGTGGCAGAGACTCTCCATGGACCGGTACCTGAACACTACTTCCGCC 2684  
QY 852 SerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsn 871  
D5 2685 ACCAAAGTGTATCCCAAGCGGTGATCTTAGACTCTGGGAAGCTCGGACGAGGATGAT 2744  
QY 872 GlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeu 891  
D5 2745 GGGGAGCTCAACAGCTTGGCCAGTGTCTTGGAGGAGATGGGCAAGAGTGAGTGTGGTA 2804  
QY 892 PheThrValSerGluAlaGluCys 899  
D5 2805 GCCATGACCACTGATGGCGATTGC 2828

## RESULT 6

US-09-306-902A-3  
; Sequence 3, Application US/09306902A  
; Patent No. 6277585  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsay  
; Masu, Masayuki  
; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 268 BUSH STREET, SUITE 3200  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA: US/09/306, 902A  
 ; APPLICATION NUMBER: US/09/306, 902A  
 ; FILING DATE: 07-May-1999  
 ; CLASSIFICATION: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OSMAN, RICHARD A  
 ; REGISTRATION NUMBER: 36,627  
 ; REFERENCE/DOCKET NUMBER: UC96-217  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 343-4341  
 ; TELEFAX: (415) 343-4342  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2831 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-09-306-902A-3

## Alignment Scores:

Pred. No.:	5,72e-229	Length:	2831
Score:	2463.00	Matches:	495
Percent Similarity:	67.41%	Conservative:	144
Best Local Similarity:	52.22%	Mismatches:	227
Query Match:	51.45%	Indels:	83
DB:	3	Gaps:	19

US-09-970-944-2 (1-899) x US-09-306-902A-3 (1-2831)

QY	9	ProAlaLeuLeuGluGlyValLeuAlaTrpLeuArgGlySerGlyAlaGlnSer	28
DB	61	CGAGCTTAGCAGGCAATT-----GACTCTGGTGCCAG-----	93
QY	29	AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuProHisPheLeuVal	48
DB	94	---GGACTCCAGACTCTCCATCCATCAGACCCCGGAGCAGCTGCCTCATTCTCTGCTG	150
QY	49	GluproGluAspValTyrIleValLysAsnLysProValLeuValCysLysAlaVal	68
DB	151	GAACCCAGAGGATGCTACATCGTAAAGAACAGCCAGTGAATTGCACTGCCGACCTTC	210
QY	69	ProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArgGlnValAspHisVal	88
DB	211	CCTGCCACACAGATCTACTTCAAGTGTAATGCGAGTGGGTAGCCAGAAAGCCACCGTC	270
QY	89	IleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnVal	108
DB	271	ACGAGGAGAGCTCGATGAGCCAGCCTTCGGAATACGAGAGGTGCAGATGAGGTG	330
QY	109	SerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAla	128
DB	331	TCGCGGACAGAGTGGAGAACTTTTGGGCTCGAGGACTACTGGTGTCTGCTGGCC	390
QY	129	TrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArg	148
DB	391	TGGAGCTCTTCGGGAACACCAAGAGTCCCGAGCTACATCCGCAATTGCCTACTTGGC	450

QY	149	LysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeu	168
DB	451	AAGAACTTTGACGAGGACCTCTGGCGAAGAGGTTACCTTGGATCATGAGTCTCTTCTG	510
QY	169	ProCysArgProProGluGlyIleProProAlaGluValGluTrpLeuArgAsnGluAsp	188
DB	511	CAGTGGCCGCCACACAGAGGAGTGCTGTGGCTGAGGTGGAATGCTCAAGAATCAAGAT	570
QY	189	LeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValVal	208
DB	571	GTTCATCGATCCCGCTCAGGACACTAACTTCCTGCTCACCATTGACCAACATCATCATC	630
QY	209	ArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAla	228
DB	631	CGCAGGCGCGCTCTCAGACACAGCCAACTACACTGTGTGGCAAGAATATTGTGGCC	690
QY	229	ArgArgSerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrp	248
DB	691	AAGCGCGGAGCAGCAGCGCCACAGTCATCTCTATGTGAACGGAGGTTGGTCCAGCTGG	750
QY	249	ThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSer	268
DB	751	GCAGAAATGTCACCTCTCTAAACCGCTGCGGCGAGGTTGCGAAGAACTACTAGGACC	810
QY	269	CysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAsp	288
DB	811	TGCACCAACCCAGCCCCACTCAATGGAGGTGCTCTTCGAGGGACAG---GCTTGCACG	867
QY	289	ArgThrValSerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrp	308
DB	868	AAGACGGCTTGCCACCCGCTGTCGCCAGTGGATGGAGCGGTGGACTGAGTGGAGCAAGTG	927
QY	309	SerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAspProAlaPro	328
DB	928	TCCGCTGCAGCACAGAGTGTGCGCACTGGCGAGCGCGAGTGTGATGCGACCCCGCCG	987
QY	329	ArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAsp	348
DB	988	CAGAACGGAGCGGTGACTGACGGGACCGTACTTGTGATCCAAAGAACTGCACCGATGGG	1047
QY	349	LeuCysVal-----HisSerAlaSerGlyProGlu-----	358
DB	1048	CTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCCTAAAGCGCCCTCGGAGCGCTCG	1107
QY	359	---AspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuValLeuLeu	376
DB	1108	GGAGACGTGGGCGCTGTATCGGGGCGCTCGTGGTGGCGCTCTTTGGTCTTCGACATTCTC	1167
QY	377	LeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAla	396
DB	1168	ATGGCTGTAGAGTGATCGTGTACCGAGAAACTGCCGGGACTTCGACGAGACATCACT	1227
QY	397	AspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAla	415
DB	1228	GACTCTCTGTGCGCTCACTGGTGTTCACCCCGCTCAACTTCAAGACTGCAAGGCC	1287
QY	416	AspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr-----Thr	432
DB	1288	AGCAACCCACAGCTCTCGACCCATCCGCCCCCTCGGACCTTAACGCGGACGCTGGCGATC	1347
QY	433	TyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsn	452
DB	1348	TACCGCGAGCTGTGTATGCGCTCGAGGACTCT---GCCGACAGATCCCTATGACTAAT	1404
QY	453	GlyHisLeuLeuSerProLeu-----	459
DB	1405	TCACCCCTTCTGGATCCCTTGCACCGCTCAAGATCAAGTCTATGACTCCAGCACCATC	1464
QY	460	-----GlyGlyGly-----	462
DB	1465	GGCTCTGGGGTGGCGCTGGCTGATGAGCGGACCTGCTGGGTGTCTTACCACCCCGGTACA	1524
QY	463	-----ArgHisThrLeuHisSerSerProThrSerGlu	474

Db 1525 TACCAGGGGATTTCTCCGGGACACCCACTTCTGCACCTCGGCGAC----- 1572  
 QY 475 AlaGluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGly 494  
 Db 1573 -----GCCAGCCTTGGTCCAG---CACCTCTCGGCGCTCCCTCGAGAC 1614  
 QY 495 ThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetLeuProAsn 514  
 Db 1615 CCAGCAGCAGGTGTCAGTGGCACCTTTGTTGCTGGGTGGAGGTGACCATTCCTCCGCG 1674  
 QY 515 ThrGlyIleSerLeuLeuLeuProProAspAlaIleProArgGlyLysIleTyrGluIle 534  
 Db 1675 ACAGGGGTGAGCTGTTGTGTCACCAATGAGCCTATCCCGGGGCAAGTTCTATGACTG 1734  
 QY 535 TyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThr 553  
 Db 1735 TATCTACGTATCAACAAGACTGAAAGACCC---CTCCACTTTCCGAAGGTTCCACAGCA 1791  
 QY 554 LeuLeuSerProIleValSerCysGlyProGlyProGlyValLeuLeuThrArgProValIle 573  
 Db 1792 GTATTGAGCCCTCGGTGACCTGCGGGCCACGGGCTCTCTGTGCGCGCTGTGTTC 1851  
 QY 574 LeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLys 593  
 Db 1852 CTCACGTGCCCCACTGCTGAGTCATTGCGGAGACTGGATCTTCAGCTCAAGACC 1911  
 QY 594 GlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluAlaProSer 613  
 Db 1912 CAGGGCCATCAGGGCCACTGGGAG---GAGTGGTGAATTTGGATGAGGAGACTCTGAAC 1968  
 QY 614 HisLeuTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGly 633  
 Db 1969 ACCCTCTGCTACTGCGAGTACAGGCTAAATCTGCCACATCTGTGTGACGAGCTGGT 2028  
 QY 634 ArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeu 653  
 Db 2029 ACCTAGCTGTTCCAGGGCGAGTCTACTTCCCGCTCCGAGTCAAGGGCTCCAGTAGCC 2088  
 QY 654 LeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHis 673  
 Db 2089 ATCTTCGCCCCAGCCCTCTGCACCTCCCTGGAGTATAGTCTCAGGGTCTACTGTCTGGAG 2148  
 QY 674 AspThrHisAspAlaLeuLysGluValValGlnLeuGlyLysGlnLeuGlyGlnLeu 693  
 Db 2149 GACACTCTCGAGACTGAAGAGGTCCTAGAGCTGGAGGAGCTCTGGGTGGCTACTTG 2208  
 QY 694 IleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIle 713  
 Db 2209 GTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTACCAACACTACGCTC-TCCTC 2267  
 QY 714 HisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluLeuPro 733  
 Db 2268 CATGACATCCCCCATGCCACTGGAGGACAACTACTGGCAAGTACCAGGAGATTCCC 2327  
 QY 734 PheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArg 753  
 Db 2328 TTCTACCATGTGGNACGGGAGCCAGAAAGCCCTGCAGTGCATCTTACCTTGAGAGA 2387  
 QY 754 ValSerProSerThrSerAspLeuAlaCysLysLeuLeuTrpValTrpGlnValGluGlyAsp 773  
 Db 2388 CATAGCTAGCTCCACTGAGTTACCTGTAAGTCTGGTGGCGGAGTAGAAGGGAA 2447  
 QY 774 GlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeu 793  
 Db 2448 GCCCAGATTTTCAGCTGCACACCGAGCTGGGT---GAGACCCCTGCTGGCTCCCTGGAT 2504  
 QY 794 AlaLeuGluSerGluAlaGlyValProAlaLeu-----ValGlyProSerAlaPheLys 811  
 Db 2505 GCATCTGCTGCTGCCCCCTGGCAATGCTGCCACACACAGCTGGACCTATGCTTCAAG 2564  
 QY 812 IleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArgGly 831

Db 2565 ATACCACGTGTCATCCGCGACAGAGATCTGCAACAGCCTGGACGCCCACTCACGGGGC 2624  
 QY 832 AlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAla 851  
 Db 2625 AATGACTGGCGGCTGTGGCAGACAGAGCTCTCCAGTGGACCGGTACCTGAACCTACTTCGCC 2684  
 QY 852 SerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsn 871  
 Db 2685 ACCAAGCTGATGCTCCACAGGCGTGTCTTAGACCTCTGGAGAGTCTGGCAGCAGGATGAT 2744  
 QY 872 GlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeu 891  
 Db 2745 GGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAAGTGTGCTG 2804  
 QY 892 PheThrValSerGluAlaGluCys 899  
 Db 2805 GCCATGACCACTGATGGCGATTGC 2828  
 RESULT 7  
 US-09-833-381-1806  
 ; Sequence 1806, Application US/09833381  
 ; Patent No. 6672186  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robison, Keith E.  
 ; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
 ; FILE REFERENCE: 5800-119  
 ; CURRENT APPLICATION NUMBER: US/09/833,381  
 ; CURRENT FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: 09/516,448  
 ; PRIOR FILING DATE: 2000-02-29  
 ; NUMBER OF SEQ ID NOS: 2050  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1806  
 ; LENGTH: 1282  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(1282)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-833-381-1806  
 Alignment Scores:  
 Pred. No.: 2,49e-188 Length: 1282  
 Score: 2039.00 Matches: 410  
 Percent Similarity: 95.58% Conservative: 1  
 Best Local Similarity: 95.35% Mismatches: 11  
 Query Match: 42.59% Indels: 9  
 DB: 4 Gaps: 2  
 US-09-970-944-2 (1-899) x US-09-833-381-1806 (1-1282)  
 QY 476 GluGluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThr 495  
 Db 3 GAGGAGTTTCGTCTCCGCGCTCTCCACCAGAACTACTTCCGCTCCCTGCCCGAGGACCC 62  
 QY 496 SerAsnMetThrTyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThr 515  
 Db 63 AGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATACA 122  
 QY 516 GlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyr 535  
 Db 123 GGAATCAGCTTCTCATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTAC 182  
 QY 536 LeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeu 555  
 Db 183 CTCAGCTGCAACAGCCGAAAGACGTGAGGTTGCCCTAGCTGGTGTTCAGACCCCTGCTG 242  
 QY 556 SerProIleValSerCysGlyProPro-GlyValLeuLeuThrArgProValIleLeu-A 575  
 Db 243 AGTCCCATCGTTAGCTGTGGACCCCTCGGCGCTCTGTCTACCCGCCAGTCACTCTGGG 302  
 QY 575 laMetAspHisCysGlyGluProSerProAspSerTrp-SerLeuArgLeuLysLysGln 594

```
|||||
Db 303 GTATGGACCATGTGGGAGCCAGCCCTGACAGCTGGAGGCTCGGCTCAAAAGCAG 362
Qy 595 SerCysGluGlySerTrpGluGlnAspVal-LeuHisLeuGlyGluGluAlaProSerHi 614
Db 363 TCGTCGAGGGGAGCTGGGAG--GATGTGCTTGACCTGGGCGAGGAGGGCCCTCCCA 419
Qy 614 sLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyAr 634
Db 420 CCTCTACTACTGCCAGCTGGAGCCAGTCGCTGCTACGCTTTCACCGAGCAGCTGAGCGG 479
Qy 634 gPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaAlaAlaAlaAlaAla 654
Db 480 CTATGCCCTGTGGGAGGGCCCTCAGCGTGCTGCGGCCAAGCGCCTCAAGCTGCTTCT 539
Qy 654 uPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAs 674
Db 540 GTTTGGCGCGGTGGCTGCACTCGCTGCACTCCAGTACACATACCTGCTGCTGCTGCTGCA 599
Qy 674 pThrHisAlaLeuLysGluValValGlnLeuGlnLysGlnLeuGlyGlyGlnLeuIl 694
Db 600 CACTCACCATGCACTCAACGTAGTGGTGAGCTGGAGAGCAGCTGGGGGGACAGCTGAT 659
Qy 694 eGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuAlaGluLeuSerIleHi 714
Db 660 CCAGGAGCCAGGGTCTGCACTTCCAGGACAGTTACACACCTTGGCCTATCCATCCA 719
Qy 714 sAspValProSerSerLeuTrpLysSerLysLeuValSerTyrGlnGluIleProPh 734
Db 720 CGATGTGCCAGCTCCCTGTGGAAGATGAAGTCTCTGTGTCAGTACCAGGAGATCCCTT 779
Qy 734 eTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgVa 754
Db 780 TTATCACATCTGGAAATGGACGACGCGTACTTGCACTGCACCTTACCCCTGGAGCGTGT 839
Qy 754 lSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGl 774
Db 840 CAGCCCCAGCACTAGTACCTGGCTTCAAGCTGTGGTGAGTGAGGGGCGAGCG 899
Qy 774 yGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAl 794
Db 900 GCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAAGGTTTGCTGAGCTGCTGC 959
Qy 794 aLeuGluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPh 814
Db 960 TCTGAGAGTAGAAGCGNGGTCCTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTT 1019
Qy 814 eLeuIleArgGlnLysIleIle-SerSerLeuAspProProCysArgArgGlyAlaAspT 834
Db 1020 CCTCATTCGGCAGAGATAATTTTCCAGCTGGACCCACCTGTAGGGGGGTGCGACT 1079
Qy 834 rpArg-ThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys 853
Db 1080 GCGGAACCTCTGGCCAGAAAATCCACCTGGACAGCCATCTCAGCTTCTTTGCCCTCCAAG 1139
Qy 854 ProSerProThrAlaMerIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn 873
Db 1140 CCCAGCCCCACAGCCATGATCTCAACCTGTGGAGGGCGGCACTTCCCCAACGGCAAC 1199
Qy 874 LeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr 893
Db 1200 CTCAGCCAGCTGGT-GCAGCAGTGGCTGGACTGGGAGC--AGACGCTGGCTCTTCCA 1255
Qy 894 ValSerGluAlaGluCys 899
Db 1256 GTGTGGAGGCTGAGTGC 1273
```

## RESULT 8

```
US-09-833-381-1807
; Sequence 1807, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
```

```
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1807
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1605)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1807

Alignment Scores:
Pred. No.: 4,72e-99 Length: 1605
Score: 1123.00 Matches: 254
Percent Similarity: 62.38% Conservative: 81
Best local Similarity: 47.30% Mismatches: 155
Query Match: 23.46% Indels: 47
Dbs: 4 Gaps: 12

US-09-970-944-2 (1-899) x US-09-833-381-1807 (1-1605)
Qy 385 CysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---IleLeuThrSer 403
Db 19 TGCGGT-----GACTTCGACACAGACATCACTGACTCATCTGCTGCTGACTGCTGACT 69
Qy 404 GlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu----- 421
Db 70 GGTTCACACCCCGTCAACTTTAAGACGCGGACCAACCCGACGCTCCTACACCCC 129
Qy 422 ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeuCysProArg 440
Db 130 TCTGTGCTCTCTGACTGACCGCAGCGCGCGGCACTACCGCGGACCCCGTGTATGCCCTG 189
Qy 441 GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly 460
Db 190 CAGGACTCC---ACGCACAAAATCCCATGACCACTCTCTCTCTGAGACCCCTTACCC 246
Qy 461 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAla----- 475
Db 247 AGCCTTAAGTCAAGGTCCACACCTCCAGCACACCGGGCTCTGGGCCAGGCTGCGCAGAT 306
Qy 475 ----- 475
Db 307 GGGGCTGACCTGTGGGGGTCTTGGCGGCTGGCACAATACCTAGCGATTTCACCGGGGAC 366
Qy 476 GluGluPheValSerArgLeuSerThrGlnAsnTyrPhe-Arg-----SerLeuPr 492
Db 367 ACCCACTTCTGCACTCGCGCAGCGCGCAGCTCGGTTCCCAGCAGCTCTTGGGCGCTGCC 426
Qy 492 oArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIl 512
Db 427 CCGAGACCCAGGAGCAGCGTCAAGCGGACCTTTGGCTGCTGGGTGGGAGCTCANCAT 486
Qy 512 eProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTy 532
Db 487 CCCGGCAGGAGGTGAGCTTGTGTCGCCCAATGGAGCCATTCCCGAGGCAAGTTCTTA 546
Qy 532 rGlu-IleTyr-LeuThrLeuHisLysProGluAspValArgLeuProLeuAla---Gly 550
Db 547 CGAAGATGATTTACTCATCAACAGGAGCAAGAGTACC---CTCCCGCTTTCAGAGGG 603
Qy 551 CysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValLeuLeuThrArg 570
Db 604 ACCCANACAGTATTGAGCCCTCGTGACCTGTGGACCCACAGGCGCTCCTGTGTGCGCG 663
Qy 571 ProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArg 590
```

```
Db 664 CCGGTCATCTCACCATGCCCACTGTGGCGAAGTCAGTGGCCGTCAGTGGATCTTTCAG 723
QY LeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGlu 610
Db 724 CTCAGACCCAGGCCCCACAGGCCACTGG--GANGAGGTGTGTGACCCCTGGATGAGGAG 780
QY AlaProSerHisLeuTyrTrpCysGlnLeuGluAlaSerAlaCysTyrValPheThrGlu 630
Db 781 ACCCTGAACACACCCCTGTACTGTCCAGCTGGAGCCAGGCCCTGTCCATCTCTGTCTGGAC 840
QY GlnLeuGlyArgPhe-AlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaAla 650
Db 841 CAGCTGGGACCTTTCGCGTGTTCACGGCGAGTCTTATCCGCTCAGCAGTCAAGCGGT 900
QY ubLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnLeuVal 670
Db 901 CCAGCTGGCGGTCTTCGCCCCCGCCCTCTGCACCTCCCTGGAGTACAGCTCCGGGTCTA 960
QY rCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuLysGlnLeuG 590
Db 961 CTGCTGGAGGACACCCCTGTAGCACTGAAGGAGGTCTGGAGCTGGAGCTGGAGCTGG 1020
QY yGlyGlnLeuLeuGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeuAr 710
Db 1021 CGGATACTTGTGGAGAGCCGAACCGCTATGTTCAAGGACAGTACCACACCTGG 1080
QY gLeuSerLeuHisAspValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrG 730
Db 1081 CCTCTCCCTCCATGACCTCCCGCCATGCGGAGGAGCAAGCTCTGCGCCCAATACCA 1140
QY nGluLeuProPheTyrHisLeuTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheTh 750
Db 1141 GGAGATCCCTCTTCTATCATTTGGAGTGGAGCCAGAGGCCCTCCATCTGCACTTTTAC 1200
QY rLeuGluArg-ValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnV 770
Db 1201 CTTGGAGAGGACAGGCTGGCTCCACAGAGCTCACTGCAAGATCTGCTGGCGCAAG 1260
QY aGluGlyAspGlyGlnSerPheSerIleAsnPheAsnLeuThrLysAspThrArgPheA 790
Db 1261 TGAAGGGAGGAGGCCAGATATTCAGCTGCATACCACTCTGGCA--GAGACACCTGCTG 1317
QY lacLeuLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyProS 808
Db 1318 GCTCCTCGACACTCTGCTCTGCCCCCTGGCGAGCACTGTACCCACCGAGCTGGACCTT 1377
QY exAlaPheLysIleProPheLeuLeuArgGlnLysIleIleSerSerLeuAspProProC 828
Db 1378 ATGCTTCAAGATCCCACTGTCCATCCGCGAGAGATATGCAACAGCTAGATGCCCCCA 1437
QY ysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuS 848
Db 1438 ACTCAGCGGGCAATGACTGGCGATGTGTAGCAGAAAGCTCTATGGACCGGTACTCTGA 1497
QY exPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgH 868
Db 1498 ATTACTTTGCCCAAGAGGAGCCCGCCCGGNTGTATCTTGACCTCTGGAGAGCTCTGC 1557
QY isPheProAsnGlyAsnLeuSerGlnLeuAlaAlaVal 881
Db 1558 AGCAGGACGATGGGACCTCAACAGGCTGNCGAGTGCCCTTG 1598
```

## RESULT 9

```
US-08-253-155A-17/c
; Sequence 17, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenio
; APPLICANT: Diaretta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/253,155A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-253-155A-17
; Alignment Scores:
; Pred. No.: 8,7e-49 Length: 771
; Score: 600.50 Matches: 131
; Percent Similarity: 67.47% Conservative: 37
; Best Local Similarity: 52.61% Mismatches: 77
; Query Match: 12.54% Indels: 6
; DB: 2
```

```
US-09-970-944-2 (1-899) x US-08-253-155A-17 (1-771)
QY 541 ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer 560
Db 769 CTTATGATGTGCCA---GATTATGCTCTCCGAGACAGTATTGAGCCCTCGGTGACC 713
QY 561 CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly 580
Db 712 TGTGACCCACAGGCTCTGCTGTGCGCCGCGTCATCTCACCATGCCCCACTGTGCC 653
QY 581 GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp 600
Db 652 GAAGTCAGTGGCCCGTGTGATCTTTTTCAGCTCAAGACCCAGGCCCCAGGCCCACTGG 593
QY 601 GluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTrpCysGlnLeu 620
Db 592 GAG---GAGTGTGTGACCTGTGATGAGGAGACCTTCAACACACCTGTACTGCCAGCTG 536
QY 621 GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu 640
Db 535 GAGCCCGAGGCTGTTCATCTCTGTGGACCACTGTGGCAGCTGGCAGCTGTTCACGGGCGAG 476
QY 641 AlaLeuSerValAlaAlaAlaLysArgLeuLeuLeuLeuPheAlaProValAlaCys 660
Db 475 TCCTATTTCCGCTCAGAGCTCAAGCGGCTCCAGTGGCGCTTTCGCCCCCGCCCTCTGC 416
QY 661 ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys 680
Db 415 ACCTCCCTGGAGTACAGACTCCGGGTCTACTGTCTGGAGGACAGGCGCTGTAGCACTGAAG 356
QY 681 GluValValGlnLeuGluLysGlnLeuGlyGlnLeuLeuLeuGlnLeuProArgValLeu 700
Db 355 GAGGTGTGAGCTGGAGCGGACTCTGGCGGATACTTGGTGGAGGAGCGGAAACCGCTA 296
```



APPLICANT: Masu, Masayuki  
APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/808,982  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UC96-217  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-808-982-4

Alignment Scores:  
Pred. No.: 1.08e-15 Length: 305  
Score: 253.00 Matches: 56  
Percent Similarity: 71.43% Conservative: 14  
Best Local Similarity: 57.14% Mismatches: 28  
Query Match: 5.29% Indels: 2  
DB: 2 Gaps: 0

US-09-970-944-2 (1-899) x US-08-808-982-4 (1-305)

QY 612 ProSerHisLeuTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGln 631  
Db 14 CCGAACACACCCCTGCTACTGAGCTGGAGCCCGGCTGTAC-ATCCTGTGGACCAG 72  
QY 632 LeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaArgLeuLys 651  
Db 73 CTGGGCACCTAGCTTTTTCAGGGCGAGTCTTATCCCGCTCAGCAGTCAAGCGGCTCCAG 132  
QY 652 LeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCys 671  
Db 133 CTGGCCCGT-TTGGCCCGCCCTCTGACCTCCCTGGAGTACAGCTCCGGGTCTACTGC 191  
QY 672 LeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGly 691  
Db 192 CTGGAGGACCGCTGTAGACTGAAGAGGTGCTGGAGCTGGAGCTGGAGCTCTGGGGCGGA 251  
QY 692 GlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709  
Db 252 TACTTGGTGGAGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCAACACCTT 305

RESULT 12

US-09-306-902A-4  
Sequence 4, Application US/09306902A  
Patent No. 6277585  
GENERAL INFORMATION:  
APPLICANT: Tessier-Lavigne, Marc  
Leonardo, E. David

Hink, Lindsay  
Masu, Masayuki  
Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/306,902A  
FILING DATE: 07-May-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UC96-217  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-306-902A-4

Alignment Scores:  
Pred. No.: 1.08e-15 Length: 305  
Score: 253.00 Matches: 56  
Percent Similarity: 71.43% Conservative: 14  
Best Local Similarity: 57.14% Mismatches: 28  
Query Match: 5.29% Indels: 2  
DB: 3 Gaps: 0

US-09-970-944-2 (1-899) x US-09-306-902A-4 (1-305)

QY 612 ProSerHisLeuTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGln 631  
Db 14 CCGAACACACCCCTGCTACTGAGCTGGAGCCCGGCTGTAC-ATCCTGTGGACCAG 72  
QY 632 LeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaAlaAlaArgLeuLys 651  
Db 73 CTGGGCACCTAGCTTTTTCAGGGCGAGTCTTATCCCGCTCAGCAGTCAAGCGGCTCCAG 132  
QY 652 LeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCys 671  
Db 133 CTGGCCCGT-TTGGCCCGCCCTCTGACCTCCCTGGAGTACAGCTCCGGGTCTACTGC 191  
QY 672 LeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGly 691  
Db 192 CTGGAGGACCGCTGTAGACTGAAGAGGTGCTGGAGCTGGAGCTGGAGCTCTGGGGCGGA 251  
QY 692 GlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709  
Db 252 TACTTGGTGGAGAGCCGAAACCGCTAATGTTCAAGGACAGTTACCAACACCTT 305

RESULT 13

US-08-985-526-4  
Sequence 4, Application US/08985526  
Patent No. 6080728  
GENERAL INFORMATION:



; APPLICANT: Mixson, James A  
 ; TITLE OF INVENTION: CARRIER:DNA COMPLEXES CONTAINING DNA  
 ; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE  
 ; TITLE OF INVENTION: THERAPY  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Connolly, Bove, Lodge, & Hutzel  
 ; STREET: 1220 Market Street, P.O. Box 2207  
 ; CITY: Wilmington  
 ; STATE: Delaware  
 ; COUNTRY: U.S.A.  
 ; ZIP: 19899  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/985,526  
 ; FILING DATE:

; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/608,845  
 ; FILING DATE: 16-JUL-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McMorrow Jr., Robert G  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (302) 658-9141  
 ; TELEFAX: (302) 658-5613  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1326 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-985-526-4

Alignment Scores:  
 Pred. No.: 2,828-14 Length: 1326  
 Score: 250.00 Matches: 104  
 Percent Similarity: 35.81% Conservative: 50  
 Best Local Similarity: 24.19% Mismatches: 152  
 Query Match: 5.22% Indels: 126  
 DB: 3 Gaps: 17

US-09-970-944-2 (1-899) x US-08-985-526-4 (1-1326)

QY 13 GlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSerAlaThrValAla 32  
 DB 92 GGACTGTTGATAGTGCA-----CTGAGTGTCTACTGTCTAGA 127  
 QY 33 AsnProValProGlyAlaAsn-----ProAspLeuLeuProHis 45  
 DB 128 ACTCAGTTTACCATTCTGCAAAAGGTCTCTGCCCATCATGCTCTCCAATGCCACAG 187  
 QY 46 PheLeuValGlu-----49  
 DB 188 TTCCTGATGGAGATGCTGCTCTCGCTGTGGCCCGAGCTCTGCGGACGATGGTGTGT 247  
 QY 50 -----ProGluAspValTyrIleValLysAsnLysProValLeuValCysLys 66  
 DB 248 CTCACATGGTCGAGTGGAGCTCTGTCTACGAGCTGTGGCAATGAATTCAGACGGCG 307  
 QY 67 AlaValProAlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArg---GlnVal 85  
 DB 308 GCGGCTCTCTGGCAT-AGCCTCAACACCGATGTGAGGGCTCTCTCGGTCCAGACACGACC 366  
 QY 86 AspHisValIleGluArgSerThrAspGlySerGlyGluProThrMetGluValArg 105  
 DB 367 TGCCACATT-----375  
 QY 106 IleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrpCysGln 125

DB 376 -----CAGGAGTGTGCACAAAAGATTATAACAGGATGGTGGCTGGAGCCAC 420  
 QY 126 CysValAlaLafTrpSerSerSerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAla 145  
 DB 421 TGGTCCCGTGGTGCATCTTCTGTGACATGTGTGTGATGGTGTGATCACAAGGATC--- 477  
 QY 146 ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGluGlnGln 165  
 DB 478 CGCTCTGCAACTCTCCAGCCCCCAGATGATGATGGAAACCTCTGTAAGGCGAAGCGGG 537  
 QY 165 yIleValLeuProCysArgProProGluGly-----175  
 DB 538 GAGACAAAGCCTGCAAGAAAGACGCTGCCCATCAATGAGGCTGGGGTCTTGGTCA 597  
 QY 176 -----IleProProAlaGluValGluTrpLeuArgAsnGluAspLeuVal 190  
 DB 598 CCATGGGACATCTGTCTGTCTGTACCTGTGGAG---GAGGGGTACAGAAACGTAGTCTCTC 654  
 QY 190 LasPProSerLeuAsp-----195  
 DB 655 TGGCTGCACTCTAGAACTGACTGAAGAGAACAAAGAGTTGGCCCAATGAGTGGAGGGGCT 714  
 QY 196 -ProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAs 215  
 DB 715 CCCCTATGCT---ATCACAAAGGAGTTCAGTACAGAAATAACGAGGAATGGACTGTT-GA 770  
 QY 215 pThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAl 235  
 DB 771 TAGCTGCACCT---GAGTGTCTACTGTCTGCAACTGATCTGCAAAAAGGTGTCTCTG 827  
 QY 235 aAlaValIleValTyrValAsnGly-----243  
 DB 828 CCCCATCAATGCTCTCCATGCCACAGTTCCTGATGAGAAAGTCTGCTCGCTGTGTG 887  
 QY 244 -----GlyTrpSerThrTrpThrGluTrpSerValCysSerAl 256  
 DB 888 GCCCAGCGACTCTGCGGAGCATGGTGGTCTCCATGGTCCGAGTGGACCTCTGTTCTTAC 947  
 QY 256 sSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAs 276  
 DB 948 GAGCTGTGGCAATGGAATTCAGCAGCGGGCGCTCTCTCGCATAGC-----CTCAA 998  
 QY 276 nGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal-----Se 292  
 DB 999 CAACCGA-----TGTGAGGGCTCTCGTCCAGACACGACCTGCCCATTCAGAGTG 1052  
 QY 292 tSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGl 312  
 DB 1053 TGACAAAAGATTATAACAGGATGGTGGTGGAGCCACTGGTCCCGTGGTCTATCTGTTC 1112  
 QY 312 yLeuAspCys-----ThrHisTrpArgSerArgGluCysSerAspProAlaProAr 329  
 DB 1113 TGTGACATGTGTGTGTGTGTGATCACAAGGATCGGCTCTGCAACTCTCCAGGCCCA 1172  
 QY 329 gAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLe 349  
 DB 1173 GATGAATGGAAACCTCTGTGAAGGGCGAAGCGGGGAGACCAAGAGCTGCAAGAAAGACGC 1232  
 QY 349 uCys---ValHisSerAlaSerGlyPro 357  
 DB 1233 CTGCCCATCAATGAGGCTGGGGTCTCT 1260

RESULT 14  
 US-09-919-172-64  
 ; Sequence 64, Application US/09919172  
 ; Patent No. 6673545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Faris, Mary  
 ; APPLICANT: Turner, Christopher M.  
 ; TITLE OF INVENTION: PROSTATE CANCER MARKERS  
 ; FILE REFERENCE: PA-0036 US  
 ; CURRENT APPLICATION NUMBER: US/09/919,172  
 ; CURRENT FILING DATE: 2001-07-30

```
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 64
; LENGTH: 7231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Inbyte ID No. 6673545 1399366.20
; NAME/KEY: unsure
; LOCATION: 5601, 5609, 7107
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-172-64

Alignment Scores:
Pred. No.: 5,64e-13 Length: 7231
Score: 250.00 Matches: 48
Percent Similarity: 45.81% Conservative: 23
Best Local Similarity: 30.97% Mismatches: 66
Query Match: 5.22% Indels: 18
DB: 4 Gaps: 3

US-09-970-944-2 (1-899) x US-09-919-172-64 (1-7231)

QY 207 ValValArgGlnAlaArgLeuAlaSerThrAlaAsnTyrThrCysValAlaLysAsnIle 226
Db 1364 ATTCAGCAGCGCGCGCTCTCGGATACGCTCAACACCGATGTGAGGGCTCTCGGTC 1423
QY 227 ValAlaArg-----ArgArgSerAlaSerAlaAlaValIle 238
Db 1424 CAGACACGACCTGCCACATTCAGGAGTGTGACAGAGATTTAAACAG----- 1471
QY 239 ValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCys 258
Db 1472 -----GATGGTGGCTGGAGCCACTGCTCCCGTGCATCTTGTCTGTGACATGT 1522
QY 259 GlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGly 278
Db 1523 GGTGATGGTGTATCACAAGGATCCGGCTCTCAACTCTCCAGCCGCCAGATGAACGGG 1582
QY 279 AlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerVal 298
Db 1583 AAACCTGTGAAGCGAAGCGGGAGACCAAGCCCTCAAGAACGACGCTCCCTATC 1642
QY 299 AspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCys----- 315
Db 1643 AATGAGGCTGGGGTCTTGTGTGTCACCATGGGACATCTGTCTCACCTGTGGAGGAGG 1702
QY 316 ThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCys 335
Db 1703 GTACAGAACGTAGTCGTCTGTGACACACCCACACCCAGTGTGGAGGCAAGGACTGC 1762
QY 336 GlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCys 350
Db 1763 GTTGTGTGTGTAAACAGAAACAGATCTGCAACAAAGCAGGACTGT 1807

RESULT 15
US-08-985-526-2
; Sequence 2, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware

; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorow Jr., Robert G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 657 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-985-526-2

Alignment Scores:
Pred. No.: 2,24e-14 Length: 657
Score: 245.50 Matches: 49
Percent Similarity: 53.28% Conservative: 16
Best Local Similarity: 40.18% Mismatches: 44
Query Match: 5.13% Indels: 13
DB: 3 Gaps: 5

US-09-970-944-2 (1-899) x US-08-985-526-2 (1-657)

QY 244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263
Db 241 GCGTGGTCTCAATGTGTGAGTGGAGTGGCTCTTCTACAGCTGTGGCAATGGAATTCAG 300
QY 264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGly 283
Db 301 CAGCGCGCGCTCTCTGCGATAGC-----CTCAACACCGA-----TGTGAGGCG 345
QY 284 GlnAsnValHisAspArgThrVal-----SerSerLeuLeuValSerValAsp 299
Db 346 TCCTCGGTCCAGACACGCGACCTGCCACATTCAGGAGTGTGACAAAGATTTAAACAGGAT 405
QY 300 GlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCys-----Thr 316
Db 406 GGTGGCTGGAGCCACTGGTCCCGGTGTCATCTTCTGTGACATGTGGTGTGATGGTGTG 465
QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
Db 466 ATCAGAGATCCGGCTCTGCAACTCTCCAGCCCCCAGATGAATGGGAAACCCCTGTGAA 525
QY 337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCys---ValHisSerAlaSer 355
Db 526 GCGAAGCGCGGGAGACCAAGCCCTGCAAGAAAGACGCTGCCCATCAATGAGGCTGG 585
QY 356 GlyPro 357
Db 586 GGTCTCT 591

Search completed: October 6, 2004, 22:39:13
Job time : 210 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 6, 2004, 18:22:36 ; Search time 1039 Seconds  
(without alignments)  
4386.468 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGMLPALLGIVLAWL.....AVAGLQPDAGLFTVSEAC 899

Scoring table:

XSgapop 10.0 , Xgapext 0.5
YGapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09970944/runat\_05102004\_112011\_1685/app\_query.fasta\_1.1095  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09970944 @CEN 1.1.622 @runat 05102004\_112011\_1685  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPELOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=10 -XGAPOP=0.5  
-XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

#### ALIGNMENTS

##### RESULT 1

US-09-970-944-1  
; Sequence 1, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Herrman, John I  
; APPLICANT: Rastelli, Luca  
; TITLE OF INVENTION: Shinkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2881  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-970-944-1

1	4787	100.0	2881	11	US-09-970-944-1	Sequence 1, Appl
2	4698.5	98.2	2752	13	US-09-918-779-1	Sequence 1, Appl
3	4698.5	98.2	2752	17	US-10-624-932-1	Sequence 1, Appl
4	4545.5	95.0	2697	15	US-10-240-154-15	Sequence 15, Appl
5	4545.5	95.0	3014	10	US-09-933-261-1	Sequence 1, Appl
6	4545.5	95.0	3014	15	US-10-256-702-1	Sequence 1, Appl
7	4379.5	91.5	3580	17	US-10-311-623-13	Sequence 13, Appl
8	2694	56.3	1787	10	US-09-933-261-2	Sequence 2, Appl
9	2694	56.3	1787	15	US-10-256-702-2	Sequence 2, Appl
10	2490.5	52.0	2295	13	US-09-972-211-55	Sequence 55, Appl
11	2490.5	52.0	2295	15	US-10-096-625-55	Sequence 55, Appl
12	2484	51.9	2860	13	US-10-087-684-1	Sequence 1, Appl
13	2484	51.9	2860	13	US-10-087-684-3	Sequence 3, Appl
14	2484	51.9	2860	13	US-10-218-779-1	Sequence 1, Appl
15	2479	51.8	2860	13	US-10-218-779-3	Sequence 3, Appl
16	2479	51.8	3485	9	US-09-816-828-18	Sequence 18, Appl
17	2479	51.8	3884	13	US-10-147-493-145	Sequence 145, App
18	2479	51.8	3884	13	US-10-145-127-145	Sequence 145, App
19	2479	51.8	3884	13	US-10-160-503-145	Sequence 145, App
20	2479	51.8	3884	13	US-10-143-118-145	Sequence 145, App
21	2479	51.8	3884	13	US-10-144-993-145	Sequence 145, App
22	2479	51.8	3884	13	US-10-158-787-145	Sequence 145, App
23	2479	51.8	3884	13	US-10-140-024-145	Sequence 145, App
24	2479	51.8	3884	13	US-10-140-808-145	Sequence 145, App
25	2479	51.8	3884	13	US-10-152-405-145	Sequence 145, App
26	2479	51.8	3884	13	US-10-127-900A-145	Sequence 145, App
27	2479	51.8	3884	13	US-10-131-820A-145	Sequence 145, App
28	2479	51.8	3884	13	US-10-142-886-145	Sequence 145, App
29	2479	51.8	3884	13	US-10-146-728-145	Sequence 145, App
30	2479	51.8	3884	13	US-10-147-499-145	Sequence 145, App
31	2479	51.8	3884	13	US-10-157-798-145	Sequence 145, App
32	2479	51.8	3884	15	US-10-028-072-145	Sequence 145, App
33	2479	51.8	3884	15	US-10-121-049-145	Sequence 145, App
34	2479	51.8	3884	15	US-10-123-904-145	Sequence 145, App
35	2479	51.8	3884	15	US-10-140-470-145	Sequence 145, App
36	2479	51.8	3884	15	US-10-175-746-145	Sequence 145, App
37	2479	51.8	3884	15	US-10-176-918-145	Sequence 145, App
38	2479	51.8	3884	15	US-10-137-865-145	Sequence 145, App
39	2479	51.8	3884	15	US-10-140-474-145	Sequence 145, App
40	2479	51.8	3884	15	US-10-142-431-145	Sequence 145, App
41	2479	51.8	3884	15	US-10-143-114-145	Sequence 145, App
42	2479	51.8	3884	15		
43	2479	51.8	3884	15		
44	2479	51.8	3884	15		
45	2479	51.8	3884	15		

Alignment Scores:

Pred. No.: 0 Length: 2881  
 Score: 4787.00 Matches: 899  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0

US-09-970-944-2 (1-899) x US-09-970-944-1 (1-2881)

Qy	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu	20
Db	87	ATGGCCGTCGCGCCGCGCTGTGGCCACGCGCTCTGGGCATAGTCCCTCGCGCTGGCTC	146
Qy	21	ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
Db	147	CGCGCTCGGGTGGCCAGCAGATGCCACCGTGGCCAAACCCAGTGGCTGGTGGCAACCG	206
Qy	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro	60
Db	207	GACCTGCTCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAAGCCA	266
Qy	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	267	GTGCTGTGTGTGCAAGGCGGTGCCCGCCAGCGAGATCTTCTTCAAGTGCAACGGGGAG	326
Qy	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	327	TGGTGGCCCGAGTGGACCACTGATCGAGCGCAGCACAGCGGGAGGAGTGTGGAGCG	386
Qy	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	387	ACCATGGAGGTCCGATTAATGTCTCAAGCGAGCAGGTCTGAGAGAGTGTTCGGGGTGGAG	446
Qy	121	GluTyrTrpCysGlnCysValAlaTrpSerSerSerGlyThrLysSerGlnLysAla	140
Db	447	GAATACTGGTGGCAGTGGTGGATGGAGCTCTCGGGCACCACCAAGAGTCAAGAGGCC	506
Qy	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
Db	507	TACATCCGATAGCCAGATTGGCAAGAACTTCGAGCAGGAGCGCTGGCCCAAGAGGTG	566
Qy	161	SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180
Db	567	TCCCTGGAGCAGGGCATCGTGTGCTCCCTGCCAGCGAGGAGTCCCTCCAGCCGAG	626
Qy	181	ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
Db	627	GTGGAGTGGCTCCGAAACAGGAGCTGTGGTGGACCCGCTCCCTGGACCCCAATGTATACATC	686
Qy	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
Db	687	ACCGCGGACACAGCTGTGGTGGACAGCGCGCTTGTGTGACACGCGCAACTACACC	746
Qy	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240
Db	747	TGGTGGCCCAAGAACATCGTGGACGTCGCGCAGCGCTCCGCTGTGTCAFCGCTAC	806
Qy	241	ValAsnGlyClyTrpSerThrTrpGluTrpSerValCysSerAlaSerCysGlyArg	260
Db	807	GTGAACGGTGGTGGTGGACGTGGACCGAGTGTCTCGTCTGGAGCGCCAGCTGTGGCGCG	866
Qy	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280
Db	867	GGCTGGCAGAAACGAGCGGAGCTGCACCAACCCCGCGCTCTCAACGGGGCGCTTTC	926
Qy	281	CysGluClyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
Db	927	TGTGAGGGGAGAAATGTGCCATGACCGCACCGTCTCTCTCTGTGTCTGTGGAGCGG	986
Qy	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
Db	987	AGCTGGAGCCCGTGGAGCAAGTGTGTGGCTGTGGGCTGGACTGGACCCACTGGCGGAGC	1046

Qy	321	ArgGluCysSerAspProAlaProArgAsnGlyClyGluGluCysGlnGlyThrAspLeu	340
Db	1047	CGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGAGGAGTGCCAGGACATGACCTG	1106
Qy	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360
Db	1107	GACACCCGCAACTGTACAGTACCTCTGTGTACACAGTGTCTTGTGGCCCTGAGGACGTG	1166
Qy	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu	380
Db	1167	GCCCTCTATGTGGCCCTCATCGCGTGGCGCTGCTGCTGGTCTGCTGTGCTCTC	1226
Qy	381	IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle	400
Db	1227	ATCCTCGTTTATTCGCCGGAAGAGGAGGGCTGGACTCAGATGTGGCTGACTGCTGCATT	1286
Qy	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420
Db	1287	CTCACCCTCAGCTTCAGCCCGTCAAGCATCAAGCCCAAGCAGCAACCCCATCTG	1346
Qy	421	LeuThrIleGlnProAspLeuSerThrThrThrThrThrThrThrThrThrThrThr	440
Db	1347	CTCACCATCCAGCGGACCTCAGCACCAACCACTACCCAGGGCAGTCTCTGTCCCGG	1406
Qy	441	GlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly	460
Db	1407	CAGATGGGCCAGCCCAAGTTCAGCTCAACATGGGACCTGCTCAGCCCTCGGT	1466
Qy	461	GlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGluGluPheValSer	480
Db	1467	GGCGGCCGCCACACACACACACACAGCTCTCCACCTCTGAGGCCGAGGAGTCTCTCC	1526
Qy	481	ArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyr	500
Db	1527	CGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGACCCAGCAACATGACCTAT	1586
Qy	501	GlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeu	520
Db	1587	GGGACCTTCAACTCTCGGGGCGGCTGATGATCCCTTAATACAGGTATCAGCCCTCCTC	1646
Qy	521	IleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLys	540
Db	1647	ATCCCCCAGATGCCATACCCCGAGGAGAAATCTATGATCTACCTACCGTGCACAAG	1706
Qy	541	ProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSer	560
Db	1707	CCGGAAGACGTGAGGTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1766
Qy	561	CysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGly	580
Db	1767	TGTGACCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1826
Qy	581	GluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrp	600
Db	1827	GAGCCAGCCCTGACAGCTGGAGCTCGCTCAAAAAGCAGTCTGTCGCGAGGGCAGCTG	1886
Qy	601	GluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeu	620
Db	1887	GAGCAGGATGTCTGACCTCGGCGAGGAGGGCCCTCCACCTCTACTTACTGCCAGCTG	1946
Qy	621	GluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGlu	640
Db	1947	GAGGCCAGTGCCTGTCTACCTCTTCACCGAGCAGTGGCGGCTTTCCTGCTGGTGGAG	2006
Qy	641	AlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCys	660
Db	2007	GGCCTCAGCGTGGCTGCCCGCAAGCGCTCAAGCTGCTCTGTTTGGCGCGTGGCTGC	2066
Qy	661	ThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLys	680
Db	2067	ACCTCCTCGTACAACTCCGGGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2126

```
QY 681 GluValValGluLeuGluLysGlnLeuGlyGlnLeuLeuLeuLeuProArgValLeu 700
Db 2127 GAGTGGTGCAGCTGAGAAAGCAGCTGGGGGACAGCTGATCAGAGACCAAGGTCCTG 2186
QY 701 HisPheLysAspSerThrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeu 720
Db 2187 CACTTCAAGAGCAGTTACCAACACCTGCGCCCTATCCATCCAGGATGTGCCAGCTCCCTG 2246
QY 721 TrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGly 740
Db 2247 TGAAGAGTAAGCTCTTGTGAGCTACCAAGAGATCCCTTTTATCACATCTGGAGTGC 2306
QY 741 ThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAsp 760
Db 2307 ACGCAGCGGTACTTGCACCTGCACCTTCACTGGAGCGTGTGAGCCCGCAGCACTAGTAC 2366
QY 761 LeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsn 780
Db 2367 CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGGAGCGGCGAGAGCTTACAGCATCA 2426
QY 781 PheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly 800
Db 2427 TTCAACATCACCAAGGACACAGAGTTTGTGAGCTGTGGCTCTGGAGAGTGAAGCGGG 2486
QY 801 ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIle 820
Db 2487 GTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATA 2546
QY 821 IleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLys 840
Db 2547 ATTTCAGCCTGGAGCCCAACCTGTAGCGGGGTGGCGAGCTGGCGGACTTGGCCAGAAA 2606
QY 841 LeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetile 860
Db 2607 CTCACCTGGAGACGCACTTCAGCTTCTTGGCTTCAAGCCCGCAGCCCGCAGCCATGATC 2666
QY 861 LeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 880
Db 2667 CTCAACCTGTGGAGGCGCGCACTTCCCAAGCGCAACCTCAGCCACCTGGCTGCAGCA 2726
QY 881 ValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2727 GTGGCTGGAGTGGGCCAGCCAGAGCGCTGGCCCTCTTCAAGTGTGGAGGCTGAGTGC 2783
```

## RESULT 2

```
US-09-918-779-1
; Sequence 1, Application US/09918779
; Publication No. US20030064369A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zernusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsbrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/09/918,779
; PRIORITY FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/221,409
```

```
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,697
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/263,662
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/281,645
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-779-1

Alignment Scores:
Pred. No.: 0 Length: 2752
Score: 4698.50 Matches: 888
Percent Similarity: 98.89% Conservative: 2
Best Local Similarity: 98.67% Mismatches: 7
Query Match: 98.15% Indels: 3
DB: 13 Gaps: 3

US-09-970-944-2 (1-899) x US-09-918-779-1 (1-2752)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeu 20
Db 46 ATGCCGCTCGCGCCCGGCTGTGGCAGCGCTCTGGGCATAGTCTCGCGCTTGGCTC 105
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyValaAsnPro 40
Db 106 CGCGGCTCGGCTGCCAGCAGAGTGCCACCGTGGCCCAACCCAGTGCCTGGTGCACACCCG 165
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db 166 GACCTGCTTCCCACTTCTGGTGGAGCCGAGGATGTGTACATCGTCAAGAACCAAGCCA 225
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 226 GTGCTGCTTGTGTGCAAGGCGCTGCCCGCCAGCATCTTCTTCAAGTGCACAGGGGAG 285
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100
Db 286 TGGGTGGCCAGGTGGACCAAGTATCGAGCGGAGCAGACAGCGGAGCAGTGGGCTGCC 345
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 346 ACCATGAGGTCCGCATTATGTCTCAAGCAGCAGGTTCGAGAGGTGTTCGGGCTGGAG 405
QY 121 GluTyrTrpCysGlnCysValAlaTTrpSerSerSerGlyThrThrLysSerGlnLysAla 140
Db 406 GAATACTGTGGCCAGTGGCGGATGCGTGGGATGAGTCTCTCGGGCCACCAAGAGTCAAGGCC 465
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 466 TACATCCGATACCCAGATTGCGCAAGACTTCCAGCAGAGCGCGCTGGCCAAAGAGGTG 525
```

Qy 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180  
Db 526 TCCCTGGAGCAGGGCATCGTCTGCCCTGCCGTCCACCGAGGGCATCCCTCCAGCCGAG 585  
Qy 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
Db 586 GTGGAGTGGCTCCGGAAACGAGACCTGGTGACCGCTCCCTGGAGCCCAATGTATACATC 645  
Qy 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
Db 646 ACGCGGAGCAGACGCTGGTGGTGCGACAGGCCCGCTTGTGACACGGCCAACTACACC 705  
Qy 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr 240  
Db 706 TGGTGGCCAAAGAACATCGTGCGACGTGCGCGACGGCTCCGCTGCTCATCGCTAC 765  
Qy 241 ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg 260  
Db 766 GTGAACGGTGGGTGTCGACGTGGACCGAGTGGTCCGCTGCGAGCGCCAGCTGTGGGCGC 825  
Qy 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
Db 826 GGCTGGCAGAAACGAGCGGAGCTGCACCAACCGCGGCGCTCTCAACGGGGGCGCTTTC 885  
Qy 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
Db 886 TGTGAGGGGCGAATGTC---CAGAAACAGCGCTGCGCCACCGCTGTGCCAGTAGACGGC 942  
Qy 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
Db 943 AGCTGAGCGCGTGGAGCAAGTGTGCGCTGTGGCTGGAGTGCACCCACTGTGGGAGC 1002  
Qy 321 A-gGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340  
Db 1003 CGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTG 1062  
Qy 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360  
Db 1063 GACACCGCAACTGTACCACTGACCTCTGTGTACACAGTGTCTGGCCCTGAGGACGTG 1122  
Qy 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380  
Db 1123 GCCCTCTATGTGGGCTCATCGCGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTCCT 1182  
Qy 381 IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle 400  
Db 1183 ATCCTCGTTATTGCCGGAAGAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242  
Qy 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
Db 1243 CTCACCTCAGGCTTCAGCGCCGTCAGCATCAAGCCCGAGCAAGACACCCCACTCTG 1302  
Qy 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr 439  
Db 1303 CTCACCATCAGCCGAGCACTCAGCACCAACACCAACCACTACAGGGCAGTCTCTGCC 1362  
Qy 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459  
Db 1363 CGGCGAGATGGGCCCCAGCCCCAAGTTCCAGCTCACCAAATGGGCACTGTCTCAGCCCCCTG 1422  
Qy 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479  
Db 1423 GGTGGGCGCGCCACACTGCACACAGCTCTCCCAACCTCTCAGGCGGAGGAGTTCGTC 1482  
Qy 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499  
Db 1483 TCCGCGCTCTCCACCCAGACTCTTCGCTCCCTGCCCGGAGGCAACCAATGACC 1542  
Qy 500 TyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519  
Db 1543 TATGGGACCTTCAACTTCTCCTGGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTC 1602  
Qy 520 LeuIleProProAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539

Db 1603 CTATATCCCCCAGATGCCATACCCCGAGGAGAGATCATGAGATCTACTCTCAGCGTCGAC 1662  
Qy 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559  
Db 1663 AAGCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTAGAGCCCTGCTGAGTCCCATCGTT 1722  
Qy 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579  
Db 1723 AGCTGTGGACCCCTGGCGTCTGTACCCGGCCAGTCATCTGGCTATGGACCACTGT 1782  
Qy 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599  
Db 1783 GGGGAGCCCGCCCTGACAGCTGGAGCCTCGCCCTCAAAAGACAGTCTGTGGAGGGCAGC 1842  
Qy 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTrpCysGln 619  
Db 1843 TGGGAG---GATGTCTGCACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCGCAG 1899  
Qy 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639  
Db 1900 CTGGAGGCGCAGTCTGTCTGTCTTACCCGAGCAGCTGGCCGCTTTGCCCTGTGGGA 1959  
Qy 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659  
Db 1960 GAGGCGCTCAGCTGGCTGCCCGCAAGCGCTCAAGCTGCTTCTCTTGGCGCGGTGGCC 2019  
Qy 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679  
Db 2020 TGCACCTCCCTCGAGTACAAATCCGGGTCTACTCGCTGCATGACACCCACCATGACATC 2079  
Qy 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699  
Db 2080 AAGGAGGTGGTGCACCTGGAGAGCAGCTGGGGGGACAGCTGATCCAGAGGCCACGGGTC 2139  
Qy 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719  
Db 2140 CTGCACTTCAAGACAGATTACCAACACCTGCGCCTATCCATCCAGATGCGCCAGCTCC 2199  
Qy 720 LeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739  
Db 2200 CTGTGGAAGAGTAAGCTCCTTGTACGTACACGAGAGATCCCTTTTATCACATCTCGAAT 2259  
Qy 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759  
Db 2260 GGCACGACGCGTACTTGCATCTGCACCTTACCTTGAGCGGTGTAGCCCCCAGCACTAGT 2319  
Qy 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779  
Db 2320 GACCTGGCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGAGCGGCAGAGCTTCAGCATC 2379  
Qy 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
Db 2380 AACTTCAAGATCACCAGGACACAGGTTTGTGTAGCTGTGGCTCTGAGAGTGAAGCG 2439  
Qy 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
Db 2440 GGGGTCCCAGCCCTGTGGCCCCCAGTGCCTTCAAGATCCCCTTCTCATTCGGCAGAG 2499  
Qy 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839  
Db 2500 ATAATTTCCAGCTGCACCCACCCCTAGCGGGGTGGCGACTGTGGGGGACTCTTGGGCCAG 2559  
Qy 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859  
Db 2560 AAATCCACTGACAGCCATCTCAGCTCTTTTGGCTTCCAAAGCCCGCCAGCCAGCCATG 2619  
Qy 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879  
Db 2620 ATCCTCAACTGTGGGAGGGCGGCACTTCCCCAACGGCAACCTCAGCAGCTGGCTGCA 2679  
Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899

Db 2680 GCAGTGGCTGGAGTGGCCAGCAGCGTGGCCCTCTCACAGTGTGGAGGCTGAGTGC 2739

RESULT 3

US-10-624-932-1

; Sequence 1, Application US/10624932

; Publication NO. US20040096877A1

; GENERAL INFORMATION:

; APPLICANT: Taupier, Raymond

; APPLICANT: Padigar, Muralidhara

; APPLICANT: Rastelli, Luca

; APPLICANT: Spaderma, Steven

; APPLICANT: Shimkets, Richard

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Spytek, Kimberly

; APPLICANT: Shenoy, Suresh

; APPLICANT: Li, Li

; APPLICANT: Gusev, Vladimir

; APPLICANT: Grosse, William

; APPLICANT: Alsobrook, John

; APPLICANT: Lepley, Denise

; APPLICANT: Burgess, Catherine

; APPLICANT: Gerlach, Valerie

; APPLICANT: Ellerman, Karen

; APPLICANT: MacDougall, John

; APPLICANT: Stone, David

; APPLICANT: Smithson, Glennda

; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-074 US

; CURRENT APPLICATION NUMBER: US/10/624,932

; CURRENT FILING DATE: 2003-07-21

; PRIOR APPLICATION NUMBER: 09/918,779

; PRIOR FILING DATE: 2001-07-03

; PRIOR APPLICATION NUMBER: 60/221,409

; PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 60/222,840

; PRIOR FILING DATE: 2000-08-04

; PRIOR APPLICATION NUMBER: 60/223,752

; PRIOR FILING DATE: 2000-08-08

; PRIOR APPLICATION NUMBER: 60/223,762

; PRIOR FILING DATE: 2000-08-08

; PRIOR APPLICATION NUMBER: 60/223,770

; PRIOR FILING DATE: 2000-08-08

; PRIOR APPLICATION NUMBER: 60/223,769

; PRIOR FILING DATE: 2000-08-08

; PRIOR APPLICATION NUMBER: 60/225,146

; PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/225,392

; PRIOR FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: 60/225,470

; PRIOR FILING DATE: 2000-08-15

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2752

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-624-932-1

Alignment Scores:

Pred. No.:	0	Length:	2752
Score:	4698.50	Matches:	888
Percent Similarity:	98.89%	Conservative:	2
Best Local Similarity:	98.67%	Mismatches:	7
Query Match:	98.15%	Indels:	3
DB:	17	Gaps:	3

US-09-970-944-2 (1-899) x US-10-624-932-1 (1-2752)

Qy 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyVileValLeuAlaAlaTrpLeu 20

Db 46 ATGGCCGTCCGCGCCGCGCTGTGGCAGCGCTCTCTGGGCGATAGTCTCGCGCTTGGCTC 105

QY	21	ArgGlySerGlyAlaGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro	40
Db	106	CGCGGCTGGGTGCCAGCAGTCCACCGTGGCCAAACCCAGTGGTGGTCCACCCG	165
QY	41	AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro	60
Db	166	GACCTGCTTCCCACTTCTCTGGTGGAGCCCGAGGATGTATCATCTGTCAAGTCAACCGGGAG	225
QY	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	226	GTGCTGCTTGTGTCAAGGCGTGGCCAGCAGATCTTCTTCAAGTCAACCGGGAG	285
QY	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	286	TGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACGAGGAGTGGGTGGCC	345
QY	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	346	ACCATGGAGGTCCGCATTAAATGTCTCAAGGCAGCAGGTGAGAGGTGTTCGGGCTGGAG	405
QY	121	GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla	140
Db	406	GAATACTGTCGTCAGTGGTGGCATGGCTCTCGGCGACCAACCAAGAGTCAGAAGGCC	465
QY	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
Db	466	TACATCCGATAGCAGATTGGCAAGAACTTCGAGCAGAGCCGCTGGCCCAAGGAGGTG	525
QY	161	SerLeuGluGlnGlyVileValLeuProCysArgProProGluGlyVileProProAlaGlu	180
Db	526	TCCCTGGAGCAGGGCATCGTGTGCTGCGCTGCCCTCACCGAGGGCATCCCTCCAGCCGAG	585
QY	181	ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
Db	586	GTGGAGTGGCTCCGAGACGAGGACCTGGTGGACCGCTCCCTGGACCCCAATGTATACATC	645
QY	201	ThrArgGluHisSerLeuValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
Db	646	ACGCGGAGCAGCAGCTGGTGGCAGAGCCCGCTTGTGTGACACGCGCCCACTACACC	705
QY	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240
Db	706	TGGTGGCCCAAGAACATCGTGGCAGTCCCGCAGCGCTCCGCTGCTCATCGTCTAC	765
QY	241	ValAsnGlyGlyTrpSerThrTrpThrTrpSerValCysSerAlaSerCysGlyArg	260
Db	766	GTGAACGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC	825
QY	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280
Db	826	GGCTGGCAGAAACGAGCGGAGCTGCACCAACCGCGCCCTCTCAACGGGCGCTTTC	885
QY	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
Db	886	TGTGAGGGCGAATGTC--CAGAAAACAGCTGGCGCCACCTGTGCCAGTAGACGCG	942
QY	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
Db	943	AGCTGGAGCCGTCGAGCAAGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC	1002
QY	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu	340
Db	1003	CGTAGTGTCTTGACCCAGCACCCCGCAACCGAGGGAGGTGCCAGGGCAGCTGACCTG	1062
QY	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360
Db	1063	GACACCGCAACTGTACCACTGACCTGTGTGTACACAGTGTCTTGGGCGCTGGAGACGTG	1122
QY	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuValLeu	380
Db	1123	GCCCTCATGTGGGCTCATCGCGTGGCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGC	1182
QY	381	IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerIle	400

```
Db 1183 ATCTCTCGTTTATGTCGGAAGAGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1242
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1243 CTCACCTCAGGCTTCAGGCCCTCAGCATCAAGCCAGCAAGAGCAGACACCCCATCTG 1302
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db 1303 CTCACCATCCAGCCGACCTCAGCACCCACCACCTACCGGAGGTCTCTGTCCC 1362
QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1363 CGGCGAGGATGGGCCCCAGCCCCAAGTTCCAGCTCACCANTGGGCACCTGCTCAGCCCCCTG 1422
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479
Db 1423 GGTGGGGGGCGGCACACATGACACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC 1482
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1483 TCCCGCCCTTCCACCCAGAACTACTTCCGCTCCCTGCCCGGAGGCACCAACATGACC 1542
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1543 TATGGAGACTTCAACTTCTCGGGGGCGGCTGATGATCCCTAAATACAGGTATCAGCCTC 1602
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
Db 1603 CTCATCCCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACCTCAGCTGCAC 1662
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1663 AAGCGGAAGACGTAGGTTGCCCTTAGCTGGCTGTACAGCCCTGCTGAGTCCATCGT 1722
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1723 AGCTGTGGACCCCTCGGCTGCTGCTACCCGCGCAGTCTATCTGGCTATGACCACTGT 1782
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1783 GGGGAGCCCCAGCCCTCAGAGCTGAGCCTCGGCTCAAAAAGCAGTGTGCGAGGGCAGC 1842
QY 600 TrpGluGlnAspValLeuHisIleuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
Db 1843 TGGGAG---GATGTGCTGCACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTCCAC 1899
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1900 CTGGAGGCCAGTCCCTGCTACGTTCTACCGAGCAGCTGGGCGCCTTGGCCCTGGTGGGA 1959
QY 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db 1960 GAGGCCCTCAGCGTGGCTGGCGGCAAGCGCCTCAAGCTGCTTCTGTTGGCGGCTGGCC 2019
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 2020 TGCACTTCCCTCAGTACAAATCCGCGTCTACTGCTGATGACACCCACGATGACATC 2079
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699
Db 2080 AAGAGGTGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTTC 2139
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2140 CTGCACTTCAAGGACAGTTACCAACCTCGCCTATCCATCCACGATGTGCCAGCTCC 2199
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2200 CTGTGAAGAGTAAAGCTCTTGTGCTACGAGAGATCCCTTTTATCATCTGGAAT 2259
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
```

```
Db 2260 GGCACGCGAGGCTACTTGCATGTCACCTTACCCTTGGAGGCTGTGAGCCCGCAGCACTAGT 2319
QY 760 AspLeuAlaCysLysLeuTyrValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2320 GACCTGGCCTGCAAGCTGTGGGTGTGGAGGTGGAGGGCGACGGGAGAGCTTTCAGCATC 2379
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuSerGluAla 799
Db 2380 AACTTCAATCATCACCAAGGACCAAGGTTTGTGAGCTGTGGCTTGGAGAGTGAAGCG 2439
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIlePhePheLeuIleArgGlnLys 819
Db 2440 GGGGTCCAGCCCTGGTGGGCCCAAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAG 2499
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2500 ATAAATTTCCAGCCTGGACCCACCTGTAGCGGGGTGGCGACTGGCGGAGCTCTGGCCAG 2559
QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
Db 2560 AAATCCACCTGGACAGCCATCTCAGCTTCTTTGGCTCCAGCCCGCAGCCCGCAGCATG 2619
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
Db 2620 ATCTTCAACCTGTGGAGGGCGGCACATTTCCCAACGGCAACCTCAGCCAGCTGGCTGCA 2679
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2680 CGAGTGGCTGGACTGGCGCAGCAGACGCTGGCCCTTTCACAGTGTGCGAGGCTGAGTGC 2739

RESULT 4
US-10-240-154-15
; Sequence 15, Application US/10240154
; Publication NO. US20030175741A1
; GENERAL INFORMATION:
; APPLICANT: Cochran et al.
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
; FILE REFERENCE: CKFW-P01-006
; CURRENT APPLICATION NUMBER: US/10/240,154
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/GB01/01486
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 2697
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2697)
US-10-240-154-15

Alignment Scores:
Pred. No.: 0 Length: 2697
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: 15 Gaps: 3

US-09-970-944-2 (1-899) x US-10-240-154-15 (1-2697)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeuValLeuAlaAlaTrpLeu 20
Db 1 ATGGCGCGTCCGGCCCGGCGCTGGCCAGTGCCTGGGCATAGTCTCGCGCGCTGCGCTT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 61 CGTGGTTCGGGTGCCCGCAGCAGTGCACGGTGGCCCAATCCAGTGCCTGGTGCACACCCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
```



Db	121	GACCTGCTGCCCACTCTCTGGTAGAGCCTGAGGACGTGTACATTGTCAGAAACAAGCCG	180
Qy	61	ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu	80
Db	181	GTGTTGTTGGTGGCAAGGCTGTGCTGCACCCACAGATCTTCTCAAGTGCATGGGGAA	240
Qy	81	TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro	100
Db	241	TGGTTCGGCCAGGTCCGATCACGTAATTAACGACGACCCACAGCAGCAGCGGATGGCA	300
Qy	101	ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu	120
Db	301	ACCATGGAGTCCGATCAACGATATCGAGGACGAGTAGAGAAAGTGTGGCTGGAG	360
Qy	121	GluTyrTrpCysGlnCysValAlaTyrSerSerSerGlyThrThrLysSerGlnLysAla	140
Db	361	GAATACGTGGTCAGTGTGGCATGGAGCTCTCGGGTACCACCAAAAGTCAGAAAGCC	420
Qy	141	TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal	160
Db	421	TACATCCGATTCGCTATTGGCAAGACTTTGAGCAGAGCCACTGGCCAGGAAGTG	480
Qy	161	SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180
Db	481	TCACCTGGAGCAAGGCACTTGTACTACCTTGTGCCCCCCAGAAAGTAATCCGCCAGCTGAG	540
Qy	181	ValGluTyrLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTCTGGAGCCCTCCCTCGATCCCAATGTGTACATC	600
Qy	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220
Db	601	ACGCGGGAGCAGACCTAGTCTGTGGTCAGGCCCGCTGGCCGACACGCCCACTACACC	660
Qy	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240
Db	661	TGTGTGGCCAGAACATCGTAGCCCGTCGCCAAGCACCTCTGCAGCGGTCAATGTTTAT	720
Qy	241	ValAsnGlyGlyTyrSerThrThrThrGluTyrSerValCysSerAlaSerCysGlyArg	260
Db	721	GTGAACGGTGGTGTGCGAGTGGAGTGTGAGTGGTCCGCTGCAGCGCCACAGCTGTGGCGGT	780
Qy	261	GlyTyrGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe	280
Db	781	GGCTGGCAGAACGGAGCGGAGCTGCACCAACCCGGACCTCTCAACGGGGCCCTTC	840
Qy	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300
Db	841	TGTGAGGGGCGAATGTC---CAGAAACACGCTGGCCACTCTGTGCCAGTGGATGGG	897
Qy	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320
Db	898	AGCTGGAGTCTCGACCCAGCACCCCGCATGAGGTGAGGAGTGTGGGGTGTGACCTG	957
Qy	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu	340
Db	958	CGGAGTGTCTGACCCAGCACCCCGCATGAGGTGAGGAGTGTGGGGTGTGACCTG	1017
Qy	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360
Db	1018	GACACCCGCACTGTACCAAGTGTACCTCTGCCTGCACACCGCTCTTGTGCCGAGGACGTG	1077
Qy	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu	380
Db	1078	GCTCTCATATCGGCTTGTGCTGTGGTGTGGCTCTTCTTGTGCTGTGGTGGCCCTT	1137
Qy	381	IleLeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIle	400
Db	1138	GGACTCATTTACTGTGCAGAGAGAGGGCTGGACTCGGATGTGGCCGACTCGTCCATC	1197
Qy	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420
Db	1198	CTCACCTCGGGCTTCCAGCTGTGAGCATCAAGCCCGACAGCAAGCAGACACCCACCTG	1257
Qy	421	LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrGlnGlySerLeuCysPro	439
Db	1258	CTCACCATCCAGCCAGACCTCAGCACCACTACCACTACCACTACCACTACCACTACCACT	1317
Qy	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu	459
Db	1318	AGGAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGGTCCACCTGTCTAGCCCACTG	1377
Qy	460	GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheVal	479
Db	1378	GGGAGTGGCGCCGATACGTTGGCACCACTCACCACCTCTCAGGCTCAGGACTTCGTC	1437
Qy	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499
Db	1438	TCCCGCTCTCCACCAAACTACTTTCCTTCCCTGCCCGGGCCAGCACCAATGGCC	1497
Qy	500	TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519
Db	1498	TACGGACCTTCAACTTCTCGGGGCCCGCTGATGATCCCTAATACGGGGATCAGCCTC	1557
Qy	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis	539
Db	1558	CTCATACCCCGGATGCCATCCCGGAGGAAGATCTACGAGATCTACCTCACCTGCAC	1617
Qy	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559
Db	1618	AAGCCAGAGAGCTGAGGTGGCCCTAGCTGCTGTCCAGCCCTGTGAGTCCAGTCTGTT	1677
Qy	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
Db	1678	AGCTGTGGGCCCCAGGAGTCTCTGTCCCGGGCAGTATCCTTGCATGACCACTGT	1737
Qy	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer	599
Db	1738	GGAGAGCCAGCCCTGACAGCTGGAGTCTGCGCTCNAAGACAGTCTCTCGCGGGCAGT	1797
Qy	600	TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrCysGln	619
Db	1798	TGGGAG---GATGTCTGCACCTTGTGAGGAGTCACTTCCACCTCTACTACTGCCAG	1854
Qy	620	LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly	639
Db	1855	CTGGAGGCCGGGGCTGTCTATGTCTTCCGAGCAGCTGGGGCGCTTGTCCCTGTGAGA	1914
Qy	640	GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla	659
Db	1915	GAGGCCCTCAGCTGGCTGCCACCAAGCGCTCAGGCTCTTCTGTTTCTCCCTGGGCC	1974
Qy	660	CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu	679
Db	1975	TGTAGCTCCCTTGAGTACAACATCCGAGTGTACTGCTACAGCAGACACCCAGCCTCTC	2034
Qy	680	LysGluValValGlnLeuLysGlnLeuGlyGlnLeuIleGlnGluProArgVal	699
Db	2035	AAGAGGTGTGAGCTGGAGAGCAGTGTGTGACAGCTGTGATCCAGAGCCTCGCGTC	2094
Qy	700	LeuHisPheLysAspSerTrpHisAsnLeuArgLeuSerIleHisAspValProSerSer	719
Db	2095	CTGCACCTTCAAAGACAGTTACCAACCTACGTCTCTCCATCCAGCAGCTGCCAGCTCC	2154
Qy	720	LeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn	739
Db	2155	CTGTGGAGAGCAAGTACTTGTGAGTACCGAGGAGATCCCTTTTACCATCTCGGAAC	2214
Qy	740	GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer	759
Db	2215	GGCACCCAGAGTATCTGCATCTGCACCTTCCCTTGGAGCGCATCAACGCCAGCACCAGC	2274
Qy	760	AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle	779
Db	2275	GACCTGGCTCTCAAGGTGTGGGTGTGGCAGGTGGAGGAGATGGCAGAGCTTCAACATC	2334

QY 780 AsnPheAsnIleThrIysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799  
 Db |||||  
 2335 AACITTCACATCACTAAGACACAGGTTGTCTGAATTGTGGCTCTGGAGAGTGAAGG 2394  
 QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
 Db |||||  
 2395 GGGTCCCAAGCCCTGGTGGCCCCAGTGCCTTCAGATCCCTTCTCTATTGGCAAAAG 2454  
 QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839  
 Db |||||  
 2455 ATCATCGCAGCTGGACCAACCTTCAGCGCGGGCGCGACTGGAGAACTCTAGCCAG 2514  
 QY 840 LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859  
 Db |||||  
 2515 AAATTCACCTGACAGCCATCTTAGCTTCTTGGCCCTCAAGCCCAAGCCCTACAGCCATG 2574  
 QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879  
 Db |||||  
 2575 ATCTCAACTATGGAGGACAGGCACCTTCCCCAACGGCAACTCGGCCAGCTGGCAGCA 2634  
 QY 880 AlaValAlaGlyLeuGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
 Db |||||  
 2635 GCTGTGGCCGAGTGGGCCAACACAGATGCTGGCCTCTTCACGGTGTGGAGGCCAGTGT 2694

RESULT 5

US-09-933-261-1  
 ; Sequence 1, Application US/09933261  
 ; Publication No. US20030040046A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tessier-Lavigne, Marc  
 ; Leonardo, E. David  
 ; Hink, Lindsay  
 ; Masu, Masayuki  
 ; Kazuko, Keino-Masu  
 ; TITLE OF INVENTION: Netrin Receptors  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 268 BUSH STREET, SUITE 3200  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/933,261  
 ; FILING DATE: 20-Aug-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/808,982  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OSMAN, RICHARD A  
 ; REGISTRATION NUMBER: 36,627  
 ; REFERENCE/DOCKET NUMBER: UC96-217  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 343-4341  
 ; TELEFAX: (415) 343-4342  
 ; INFORMATION FOR SEQ ID NO. 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3014 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ; US-09-933-261-1  
 ; Alignment Scores:

Pred. No.: 0 Length: 3014  
 Score: 4545.50 Matches: 852  
 Percent Similarity: 96.78% Conservative: 19  
 Best Local Similarity: 94.67% Mismatches: 26  
 Query Match: 94.96% Indels: 3  
 DB: 10 Gaps: 3  
 US-09-970-944-2 (1-899) x US-09-933-261-1 (1-3014)  
 QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaLafTrpLeu 20  
 Db |||||  
 1 ATGGCGCTCCGGCCCGCTGTGGCCAGTGTCTCTGGCATAGTCTCTCGCCGCTGGCTT 60  
 QY 21 ArgGlySerGlyAlaGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
 Db |||||  
 61 CCGTGTGGGTGGCCAGCAGAGTGCACGGTGGCCAAATCCAGTGGCCGCTGGTCCAAACCCC 120  
 QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTrpIleValLysAsnLysPro 60  
 Db |||||  
 121 GACCTGCTGCCCCACTTCTCTGTAGAGCTGTAGGAGCTGTACATTGTCAAGAAACAAGCCG 180  
 QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
 Db |||||  
 181 GTGTGTGTGGTGTCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGTCAATGGGAA 240  
 QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluPro 100  
 Db |||||  
 241 TGGGTCCGCCAGGTGCATCACGTAATTGAACGAGCAGCAGCAGCAGCAGCGGATGGCCA 300  
 QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
 Db |||||  
 301 ACCATGGAGTCCGTATCAACGTATCGAGCAGCAGTAGAGAAAGTGTGTGGCTGGAG 360  
 QY 121 GluTrpTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLysSerGlnLysAla 140  
 Db |||||  
 361 GAATACTGTGGCAGTGTGTGGCATGGAGTCTCTCGGTACCACCAAAAATCAGAGGCC 420  
 QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160  
 Db |||||  
 421 TACATCCGATTGCCTATTTCGCAAGAACTTTGACGAGGACCATGTGGCCAAAGAGTG 480  
 QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIleProProAlaGlu 180  
 Db |||||  
 481 TCACTGGAGCAAGGCATTGTACTTCTGTGCCCCCAGAGGAATCCCCAGCTGAG 540  
 QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
 Db |||||  
 541 GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTATC 600  
 QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
 Db |||||  
 601 ACGGGGAGCAGCAGCTAGTGTGCGCTGAGCCCGCTGGCCGACAGCGGCAACTACACC 660  
 QY 221 CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaValIleValTyr 240  
 Db |||||  
 661 TGTGTGCCCAAGAACATCGTAGCCCTGCGCGAAGCACCTCTCGACGGTCAATTGTTAT 720  
 QY 241 ValAsnGlyGlyTrpSerThrTrpTrpGluTrpSerValCysSerAlaSerGlyArg 260  
 Db |||||  
 721 GTGAACGGTGGGTGGTTCGACGTGAGTGTGCTGTGACGCGCCAGCTGTGGCGGT 780  
 QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280  
 Db |||||  
 781 GGCTGGCAGNAACGGAGCGGAGCTGCACCAACCCCGCACCTCTCAACGGGGCGCCTTC 840  
 QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300  
 Db |||||  
 841 TGTGAGGGGAGAAATGTC---CAGAAAAACAGCCCTGGCCCACTCTGTGCCAGTGGATGGG 897  
 QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320  
 Db |||||  
 898 AGCTGGAGTTCGTGGAGTAAGTGGTGTGGCTGTGGCTGTGACCTGCACCCCACTGGCGGAGC 957

```
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeu 340
Db 958 CGCGAGTGCTCTGACCCAGCACCCCAATGAGGTGAGAGTGTGGGGTGTGACCTG 1017
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
Db 1018 GACACCGCAACTGTACAGTACCTCTGCTGCACACCGCTTCTTGGCCCGAGGACGTG 1077
QY 361 AlaLeuValGlyLeuLeuAlaValAlaValCysLeuValLeuLeuValLeu 380
Db 1078 GCTCTCTACATCGCCCTTGTGCTGTGGTGTGTGCTTCTTGTGTGTGGCCCTT 1137
QY 381 IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerIle 400
Db 1138 GGACTCATTTACTGTGCGAAGAAGAGGGCTGGACTCCGATGTGGCCGACTCGTCCATC 1197
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
Db 1198 CTCACCTCGGGCTTCCAGCCTGTGCAGCTCAAGCCAGCAAGACAGACAACCCCACTG 1257
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysPro 439
Db 1258 CTCACCATCCAGCCAGACCTCAGCACCACTACCACTACAGGGCAGTCTATGTTCG 1317
QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
Db 1318 AGCGAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGGTCACTGTCTAGCCCACTG 1377
QY 460 GlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479
Db 1378 GGGAGTGGCGGCATACGTGTGCACACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
Db 1438 TCCCGCCTCTCACCCAAACTACTTGTTCCTCCCGCCGCGGACCAAGCAACATGGCC 1497
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
Db 1498 TACGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACGGGGATCAGCCTC 1557
QY 520 LeuIleProAspAlaIleProArgGlyLysIleTyrGluIleTyrIleThrLeuHis 539
Db 1558 CTCATACCCCGGATGCCATCCCGGAGAAAGATCTACGAGATCTACCTCACACTGCAC 1617
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
Db 1618 AAGCCAGAAAGCTGAGGTGCCCCCTAGCTGCTGTGCAGACCTGTCTGAGTCCAGTCGTT 1677
QY 560 SerCysGlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
Db 1678 AGCTGTGGGCCCCCAGGAGTCTGCTCACCCGCGCAGTCACTCTTGCAATGGACCACTGT 1737
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1738 GGAGAGCCCGCCCTGACAGCTGGAGTGTCCGCTCAAAAGACAGTCTCTGCGAGGGCAGT 1797
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln 619
Db 1798 TGGGAG---GATGCTGACCTTGGTGGAGGAGTCACTTCCCACTCTACTACTGCCAG 1854
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGGCGGGCCCTGCTATGCTTTCAGGAGAGCTGGGCGCTTGGCCCTGGTAGGA 1914
QY 640 GluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAla 659
Db 1915 GAGGCGCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCTTCTGCTTCTCCCGTGGCC 1974
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTACGTCCTTGTAGTCAACATCGAGTGTACTGCTTACAGACCCACGACGCTCTC 2034
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal 699

2035 AAGGAGTGCTGCAGCTGGAGAGACGACTAGGTGGACAGCTGATCCAGAGCCTCCGCTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACTTCAAAAGACAGATTACCAACACCTACGTCTCTCCATCCACGACGTGCCAGCTCC 2154
QY 720 LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAGAGCAAGCTACTTGTCACTACCAAGGAGATCCCTTTTACCACATCTCGAAC 2214
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGCACTAICTGCACTGCACCTTCACTCCCTGGAGGAGATGGGACAGCTTCAACATC 2274
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCCTGCAAGGTGTGGGTGTGGCAGGTGGAGGAGATGGGACAGCTTCAACATC 2334
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla 799
Db 2335 AACTTCAACATCACTAAGGACACAAGGTTTGTGTAATTCTTGGCTCTGGAGAGTGAAGGG 2394
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCACAGCCTGTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCAATCGGCAAAAG 2454
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATCATCGCAGCTGTGACCCACCTCTGCAGCCGGGGCGGACTGGAGAACTCTAGCCCAAG 2514
QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
Db 2515 AAACCTTCACTGGACAGCCTCTTAGCTTCTTGTGCTTCCAGCCAGCCCTACAGCCATG 2574
QY 860 IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 879
Db 2575 ATCCTCAACCTATGGAGGACGCGCACTTCCCAACGGCAACCTCGGCAGCTGGCAGCA 2634
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db 2635 GCTGTGGCGGAGTGGGCCAACCAAGATGCTGGCCTCTTACCGTGTGGAGGCCGAGTGT 2694

RESULT 6
US-10-256-702-1
; Sequence 1, Application US/10256702
; Publication No. US20030059859A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/256,702
; FILING DATE: 27-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
```

```
;
; FILLING DATE: 20-Aug-2001
; APPLICATION NUMBER: 08/808,992
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3014 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-256-702-1

Alignment Scores:
Pred. No.: 0 Length: 3014
Score: 4545.50 Matches: 852
Percent Similarity: 96.78% Conservative: 19
Best Local Similarity: 94.67% Mismatches: 26
Query Match: 94.96% Indels: 3
DB: 15 Gaps: 3

US-09-970-944-2 (1-899) x US-10-256-702-1 (1-3014)
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuGlyLeuValLeuAlaLTrpLeu 20
DB 1 ATGGCCGCTCGGGCCCGGCTGTGGCCAGTCTCTGGGCATAGTCTCGCCGCTGGCT 60
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
DB 61 CGTGGTTCGGGTGCCCGAGAGTGCACGGTGGCCCAATCCAGTGCCTGGTCCCAACCC 120
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValIysAsnLysPro 60
DB 121 GACCTCTGCCCACTTCCTGTGTAGAGCCTGTAGAGCGTGTACATTTGCAAGAACCAAGCCG 180
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
DB 181 GTGTTGTGTGTGCAAGCTGTGCTGTGCACCCAGATCTTCTCAAGTGCATATGGGAA 240
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerGlyGluPro 100
DB 241 TGGGTCCGCGAGTTCGATCACGTAAATTGAACGAGCAGCACGAGCGGATTGCCA 300
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
DB 301 ACCATGGAGTCCGATATCAACGTATTCAGGCGACAGTGTAGAGAAAGTGTGGGCTGGAG 360
QY 121 GluTyrTrpCysGlnCysValAlaLTrpSerSerGlyThrThrLysSerGlnLysAla 140
DB 361 GAATATGTGTGCAGTGTGTGGCATGAGCTCTCGGGTACCCCAAAAGTCAGAGGCC 420
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
DB 421 TACATCCGATTGCCATTTCGCAAGAACTTTGAGCAGGAGCCTATGGCCCAAGGAGTG 480
QY 161 SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu 180
DB 481 TCATCGAGCAGGCAATGTACTTACCTTGTGCCCCCCCCAGAGAAATCCCCCAGCTGAG 540
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200
DB 541 GTGAGTGGCTTCGAATGAGGACCTCTGTGGACCCCTCCCTCGATCCCAATGTGTACATC 600
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaAlaArgLeuAlaAspThrAlaAsnThr 220
DB 601 ACGCGGAGCACAGCCTAGTGTGTGCTGAGGCCCGCTGCGCCGACACGCGCAACTACACC 660
```

```
QY 221 CysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaAlaValIleValTyr 240
DB 661 TGTGTGGCCCAAGAACATCGTAGCCCGTCCGCAAGACCTCTGCAGCGGTCAATTGTTAT 720
QY 241 ValAsnGlyGlyTrpSerThrTrpGluTrpSerValCysSerAlaSerCysGlyArg 260
DB 721 GTGAACGGTGGTGTTCGACGTGGACTGAGTGGTCCGTCTGCAGCCGACGCTGTGGCGCT 780
QY 261 GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPhe 280
DB 781 GGCTGGCAGAAACGGAGCGGAGCTGCACCAACCGGCACCTCTCAACCGGGCGCGCTTC 840
QY 281 CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly 300
DB 841 TGTGAGGGGCAGAAATGTC---CAGAAAACAGCGCTCGCCACTCTGTGCCAGTGGATGG 897
QY 301 SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer 320
DB 898 AGCTGGAGTTCGTGGAGTAAGTGTTCAGCTGTGGCTTGACTGCACCCACTGGCGGAGC 957
QY 321 ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu 340
DB 958 CGCGAGTGTCTGTACCCAGCACCCCGCAATGAGGTGAGGAGTGTGGGGTGTGCACCTG 1017
QY 341 AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal 360
DB 1018 GACACCGCACTGTACCACTGACCTGTCTGCACACCGCTTCTTGGCCCGGAGGACGT 1077
QY 361 AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu 380
DB 1078 GTCTCTACATCGGCTTGTCTGTGGCTGTGTCTCTTCTGTGTGTGGCCCTT 1137
QY 381 IleLeuValTyrCysArgLysGlyGluLeuAspSerAspValAlaAspSerSerIle 400
DB 1138 GCACCTACTTACTGTCTGCAAGAAAGGAGGCTGCATCCGAGTGTGGCCGACTCGTCCATC 1197
QY 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420
DB 1198 CTCACCTCGGGTTCAGCTGTCTGAGCATCAAGCCAGCAAGCAGACAAACCCCACTG 1257
QY 421 LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr 439
DB 1258 CTCACCATCCAGCCAGACCTCAGCACCACTACCACTACCACTACCACTACCTATGTCG 1317
QY 440 ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu 459
DB 1318 AGGCGAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGGTCACCTGTCTCAGCCACTG 1377
QY 460 GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal 479
DB 1378 GGGAGTGGCGCCATACGTGTGCACACAGCTCACCACTCTGAGGCTGAGGACTTCGTC 1437
QY 480 SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr 499
DB 1438 TCCCGCTCTCCACCCCAAACTACTTTCGTTCCCTGCCCGCGGCAACCAACATGGCC 1497
QY 500 TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu 519
DB 1498 TACGGGACCTTCAACTTCTCGGGGCGCGCTGATGATCCCTAATAACGGGGATCAGCCTC 1557
QY 520 LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis 539
DB 1558 CTCATACCCCGGATGCCATCCCCGAGGAAGAATCTACAGATCTACCTCACATGCAC 1617
QY 540 LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal 559
DB 1618 AAGCCAGAAAGACGTGAGGTGGCCCTAGCTGGCTGTGCAGACCTGTGAGTGCAGTCGTT 1677
QY 560 SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys 579
DB 1678 AGCTGTGGGCCCCCAGGAGTCTGTCTCACCCGCGCAGTCTATCTTCAATGGACCACTGT 1737
```

```
QY 580 GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer 599
Db 1738 GGAGAGCCCGACGCTGACAGCTGGAGTCTGGCGCTCAAAAAGCAGTCTCTCGGAGGCGAGT 1797
QY 600 TrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTrpCysGln 619
Db 1798 TGGGAG---GATGTGCTGCACCTTGGTGAGAGTCACTCTCCACCTCTACTACTGCCAG 1854
QY 620 LeuGluAlaSerAlaCysTyrValPheThrGlnGlnLeuGlyArgPheAlaLeuValGly 639
Db 1855 CTGGAGCGGGCGCTGCTATGCTTCAACGAGCAGCTGGCGCGCTTGGCCCTGGTAGGA 1914
QY 640 GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla 659
Db 1915 GAGGCCCTCAGCGTGGCTGCCACCAAGCGCTCAGGCTCCTCTGTTTGTCTCCGTGGCC 1974
QY 660 CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu 679
Db 1975 TGTACGTCTCTGAGTACAACTCCGAGTGTACTGCTTACGACACCCACGACGCTCTC 2034
QY 680 LysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgVal 699
Db 2035 AAGGAGGTGGTGCAGCTGGAGAGCAGCTAGTGGACAGTATCCAGGAGCCTCGCGTC 2094
QY 700 LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer 719
Db 2095 CTGCACCTTCAAAAGACAGTTACCACAACTACGCTCTCTCCATCCAGCAGCTGCCAGCTCC 2154
QY 720 LeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn 739
Db 2155 CTGTGGAAGAGCAAGCTACTTGTGACGTACCAAGGAGATCCCTTTTACCACATCTGGAAC 2214
QY 740 GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer 759
Db 2215 GGCACCCAGCAGTATCTGCACCTTCACTTGGAGCGCATCAACCCAGCAGCACG 2274
QY 760 AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle 779
Db 2275 GACCTGGCGCTGCAAGGTGGGTGGTGGCAGGTGGAGGAGATGGGCAGAGCTTCAACATC 2334
QY 780 AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuGluSerGluAla 799
Db 2335 AACITTCACATCACAAGGACACAGAGTTGTGTAATTGTGGCTGTGGAGAGTGAAGG 2394
QY 800 GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819
Db 2395 GGGGTCCCAGCCTCGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCAATCGGCAAAAG 2454
QY 820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
Db 2455 ATCATCGCCAGTCTGGACCCACCTCGACCGGGGCGCGACTGGAGAACTTAGCCCGAG 2514
QY 840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859
Db 2515 AAACITTCACCTGGACAGCATCTTAGCTTCTTTGCTCTCAACCCAGCCGCTACAGCCATG 2574
QY 860 IleLeuAsnLeuTrpGluAlaAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 879
Db 2575 ATCCTCAACCTATGGAGGCGCAGGCACTTCCCAACGGCAACCTCGGCAGCTGGCAGCA 2634
QY 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlyCys 899
Db 2635 GCTGTGGCGGAGTGGGCCAACAGATGCTGSCCTCTTACAGGTCTCGGAGGCCAGTGT 2694
RESULT 7
US-10-311-623-13
; Sequence 13, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLIUCK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Darniel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
```

```
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 3580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CB1
US-10-311-623-13
```

## Alignment Scores:

```
Pred. No.: 0 Length: 3580
Score: 4379.50 Matches: 838
Percent Similarity: 93.22% Conservative: 1
Best Local Similarity: 93.11% Mismatches: 2
Query Match: 91.43% Indels: 59
DB: 17 Gaps: 3
```

US-09-970-944-2 (1-899) x US-10-311-623-13 (1-3580)

```
QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaIleTrpLeu 20
Db 4 ATGGCGGTCGGCGCCGGCTGTGGCCAGCGCTCTGGGGCATAGTCTCGCCGCTTGGCTC 63
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40
Db 64 CGCGGCTCGGGTGGCCAGCAGAGTGCACCGTGGCCACCCAGTGCCTGGTGCCACCCG 123
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrIleValLysAsnLysPro 60
Db 124 GACCTGCTTCCCACTTCTCTGTGGAGCCCGAGGATGTGTACATCGTCAAGAACCAAGCA 183
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80
Db 184 GTGCTGCTGTGTGCAAGCGCGTGGCCCGCCAGCAGATGCTTCTTCAAGTGAACCGGGAG 243
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerGlyGluPro 100
Db 244 TGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACGCGGAGCAGTGGGTGGTCCC 303
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120
Db 304 ACCATGGAGGTCCGCATTTAATGTCTCAAGGCAGCAGGTTCGAGAAGTGTCTCGGGCTGGAG 363
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140
Db 364 GAATACTGGTCCAGTGGTGGCATGGAGTCCCTGGGCACCAACCAAGAGTCAGAAGGCC 423
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluVal 160
Db 424 TACATCGCATAGCTATTTCGGCAAGNACTTCGAGCAGGAGCGCGCTGCCCAAGAGGTG 483
```

QY	161	SerLeuGluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGlu	180	QY	520	LeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHis	539
DB	484	TCCTGGAGCAGGCGCATCGTGCTGCCCTGCCGTCCACGGAGGGCATCCCTCCAGCCGAG	543	DB	1393	CTACCTCCCCAGATGCGCATACCCAGAGGAGATCTATGAGATCTACTCAGCTGCAC	1452
QY	181	ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle	200	QY	540	LysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleVal	559
DB	544	GTGGAGTGGCTCCGGAACAGAGACCTGGTGGACCCGTCCTCGACCCCAATATACATC	603	DB	1453	AAGCCGGAGAGCTGAGGTGGCCCTAGCTGGCTGTGCAGACCTGTGTAGTCCCATCGTT	1512
QY	201	ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr	220	QY	560	SerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCys	579
DB	604	ACGGGGAGCAGCAGCTGTGTGTGGACAGGCCCGCTGTGTGACACGGCCAACTACACC	663	DB	1513	AGCTGTGAGACCCCTTGGGTCTCTGTCTACCCGGCCAGTCATCTGGCTATGACCACTGT	1572
QY	221	CysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyr	240	QY	580	GlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySer	599
DB	664	TGCGTGGCCAGAACATCGTGACAGTCGCCGAGCGCTCCGCTGTCTCATCGCTAC	723	DB	1573	GGGGAGCCAGCCCTGACAGCTGGAGCCTGCCCTCAAAAAGCAGCTGTGCGAGGCGCAGC	1632
QY	241	ValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArg	260	QY	600	TrpGluGlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGln	619
DB	723	-----	723	DB	1633	TGGGAG--GATGTGCTGCACCTGGCGGAGGAGGGCCCTCCACCTCTACTACTGCCAG	1689
QY	261	GlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyAlaPhe	280	QY	620	LeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGly	639
DB	723	-----	723	DB	1690	CTGGAGGCCAGTGCCTGTCTTACCGGAGCAGCTGGGCGCTTTGCCCTGGTGGGA	1749
QY	281	CysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGly	300	QY	640	GluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAla	659
DB	724	-----	732	DB	1750	GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCTCAAGCTGTCTTCTGTTCGGCGGTGGCC	1809
QY	301	SerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSer	320	QY	660	CysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeu	679
DB	733	AGCTGAGGCCCGTGGAGCAAGTGGTGGCTGTGGGCTGGACTGCACCCACTGGCGGAGC	792	DB	1810	TGCACCTCCCTCGAGTACAAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCATC	1869
QY	321	ArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeu	340	QY	680	LysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgVal	699
DB	793	CGTGTGTCTCTGACCCAGCACCCCGCAACGGAGGGGAGGTGCCAGGGCACTGACCTG	852	DB	1870	AAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGGGACAGCTGATCCAGAGGCCACGGGC	1929
QY	341	AspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspVal	360	QY	700	LeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSer	719
DB	853	GACACCGCAACTGTACCACTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG	912	DB	1930	CTGCACTTCAAGACACTTACCAACCTGGCCCTATCCATCCACGATGTGCCACGCTCC	1989
QY	361	AlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeu	380	QY	720	LeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsn	739
DB	913	GCCTCTATGTGGGCTCATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTC	972	DB	1990	CTGTGGAAGAGTAAGCTCTTGTGCTACAGGAGATCCCTTTTATCACATCTGGAAT	2049
QY	381	IleLeuValTyrCysArgLysLysGlyLeuAspSerAspValAlaAspSerSerIle	400	QY	740	GlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSer	759
DB	973	ATCCTCGTTTATTGCCGGAAGAGGGGCTGGACTCAGATGTGCTGACTCGTCCATT	1032	DB	2050	GGCACACGCGGTACTTGCACCTGCACCTTACCCCTGGAGCGGTGACGCCACGACTAGT	2109
QY	401	LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu	420	QY	760	AspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIle	779
DB	1033	CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCAGCAAGCAGACACACCCCATCTG	1092	DB	2110	GACCTGGCTTGCNAGCTGTGGGTGTGGAGGTGGAGGGCGAGCGGCGAGCTTCAGCATC	2169
QY	421	LeuThrIleGlnProAspLeuSer---ThrThrThrThrThrThrThrThrThrThr	439	QY	780	AsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAla	799
DB	1093	CTCACCATCCAGCGGACCTCAGACCAACACACACACCTACAGGGGAGTCTCTGTCCC	1152	DB	2170	AACTTCAACATCACCAAGGACACACAGGTTTGTGAGCTGTGTGCTGTGGAGGTGAAGCG	2229
QY	440	ArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeu	459	QY	800	GlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys	819
DB	1153	CGGCAGGATGGGCCAGCCCAAGTTCCAGCTCACCACCAATGGGCACTGTCTCAGCCCTCG	1212	DB	2230	GGGGTCCAGCCCTGTGGGGCCCCAGTGTCTTCAAGATCCCTTCTCATTCGGCAGAAG	2289
QY	460	GlyGlyGlyArgHisThrLeuHisHisSerSerProThrSerGluAlaGluPheVal	479	QY	820	IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln	839
DB	1213	GGTGGGGCGCCACACACTGTGACACAGCTCTCCCACTTCCCACTTCCAGGCGGAGGTTCGTC	1272	DB	2290	ATAATTTCCAGCTTGACCCACCCCTGTAGCGGGGTGCCGACTGCGGACTCTGGCCAG	2349
QY	480	SerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThr	499	QY	840	LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet	859
DB	1273	TCCGGCTCTCCACCCAGAACTACTTCGGTCCCTCCCGAGGACCAACCAATGACC	1332	DB	2350	AAACTCCACTGCAGACCACTCTCAGCTTCTTTGCTCCAAAGCCCGACCCACAGCCATG	2409
QY	500	TyrGlyThrPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeu	519	QY	860	IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	879
DB	1333	TATGGGACCTTCAACTTCTTCGGGGCGGCTGTATGATCCCTAATACAGGAATCAGCCTC	1392	DB	2410	ATCCTCAACTGTGGAGGGCGGCGACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCA	2469
QY				QY	880	AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899

```

Db      2470 GCAGTGGCTGGAGTGGCCAGCAGCTGGCTCTTACAGTGTGGAGGCTGAGTGC 2529
RESULT 8
US-09-933-261-2
; Sequence 2, Application US/09933261
; Publication No. US20030040046A1
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; Leonardo, E. David
; Hink, Lindsay
; Masu, Masayuki
; Kazuko, Keino-Masu
; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; Filing DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; Filing DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1787 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-933-261-2
Alignment Scores:
Pred. No.: 5,35e-274 Length: 1787
Score: 2694.00 Matches: 543
Percent Similarity: 97.14% Conservative: 1
Best Local Similarity: 96.96% Mismatches: 11
Query Match: 56.28% Indels: 11
DB: 10 Gaps: 2
US-09-970-944-2 (1-899) x US-09-933-261-2 (1-1787)
Qy      344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363
Db      3 AACTGTACCACTGACCTCTG-TACACACTGCTTCTGGCCCTGAGGAGCTGGCCCTCAT 61
Qy      364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuValLeuIleLeuVal 383
Db      62 GTGGGCGCTCATCGCGTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Qy      384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403
Db      122 TATTGCCGGAAGAGGAGGGGCTGACTCAGATGGCTGACTCGTCCATTCTCACCTCA 181

```

```

Qy      404 GlyPheGlnProValSerIle-LysProSerLysAlaAspAsnProHisLeuLeuThrIle 423
Db      182 GGCTTCCAGCCCGTACGATCTAAGCCCAAGCAGCAACCCCATCTGCTCCACCAT 241
Qy      423 eGlnProAspLeuSer---ThrThrThrThrTyrGlnGlySerLeuCysProArgGlnAs 442
Db      242 CCAGCGGACCTCAGACCAACCACCACTACAGGGGAGTCTCTGTGTCCTCCGCGGAGGA 301
Qy      442 pGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyG 462
Db      302 TGGGGCCAGCCCAAGTTCAGCTACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCG 361
Qy      462 ValGHisThrLeuHisHisSerSerProThrSerGluAlaGluGluPheValSerArgLe 482
Db      362 CCGCCACACACTGCACCACTCTCCACCTCTGAGGCGGAGGAGTTCGTCTCCCGCCT 421
Qy      482 uSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyTh 502
Db      422 CTCACCCAGAACTACTTCGGTCTCCCTGCCCGAGGCACAGCAACATGACCTATGGGAC 481
Qy      502 rPheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePr 522
Db      482 CTTCAACTTCTCGGGGGCGGCTGATGATCCTTAATACAGGAATCAGCTCTCATGCC 541
Qy      522 oProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProG 542
Db      542 CCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCACAAGCCGA 601
Qy      542 uAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysG 562
Db      602 AGACCTGAGGTTGCCCTAGCTGGCTGTGAGACCTCTGAGCTCCATCGTAGCTGTGG 661
Qy      562 yProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluPr 582
Db      662 ACCCGCTGGCGTCTGCTCACCGGCCAGTATCTGCTGCTATGGACCACTGTGGGGAGCC 721
Qy      582 sSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluG 602
Db      722 CAGCCCTCAGCAGCTGGAGCTGGCCCTCAAAAGCAGTCTGTCGAGGGAGC-TGGAG-- 778
Qy      602 nAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAl 622
Db      779 -GATGT-CTGCACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTCCAGCTGGAGGC 836
Qy      622 aserAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLe 642
Db      837 CAGTGCCTGCTACGCTCTTCAACCGAGCAGCTGGCGCGCTTTGCCCTGGTGGGAGGCCCT 896
Qy      642 uSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSe 662
Db      897 CAGCGTGGCTGCCGCAAGCGCTCAAGCTGCTTCTGTTGCGCGGTGGCGCTGCACCTC 956
Qy      662 rLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluVa 682
Db      957 CCTCGAGTACAACTCCGGGTCTACTGCTGATGACACCCACCATGACCTCAAGAGAGT 1016
Qy      682 lValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPh 702
Db      1017 GGTGCAGCTGCAGAAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACCTC 1076
Qy      702 eLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLy 722
Db      1077 -AAGGACAGTTACCAACACCTGCC-CTATC-ATCCACGATGTGCCAGCTCCCTGTGGAA 1133
Qy      722 sSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGl 742
Db      1134 GAGTAAGCTCTTGTGCTAGCTACAGGAGATCCCCCTTTTATCACATCTGGAATGGCAGCA 1193
Qy      742 nArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAl 762
Db      1194 GCGGTACTTGCACTGCACCTTCACCTTGGAGCGTGTGAGCCCCAGCCTAGTAGCTGGC 1253
Qy      762 aCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAs 782

```









[illegible]



Db 484 GCGGGACACCAAGAGTCCGCGAGCTACGTCCGCATCGCTGTCTGCGAAGACTTC 543  
 Qy 152 GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyLeValLeuProCysArg 171  
 Db 544 GATCAGAGGCTCTGGGCAAGAGAGTGCCTTGGACCATGAGGTTCCTCTGCAAGTCCGC 603  
 Qy 172 ProProGlnGlyLeProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAsp 191  
 Db 604 CCGCGGAGGGGTGCTGTGCGGAGGTGGATGGCTCAAGAATGAGGATGTTCATCGAC 663  
 Qy 192 ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla 211  
 Db 664 CCAACCCAGSACACCACTTCCTGCTCACCATCGACCACTCATCATCGCCAGGCC 723  
 Qy 212 ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArg 231  
 Db 724 CGCTGTGCGACACTGCCAACTATACCTGCTGGCCAAAGAACATCTGTCGCAACGCCG 783  
 Qy 232 SerAlaSerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrp 251  
 Db 784 AGCACCTGCCACCGCTCATCTGCTACGTGAATGGCGGCTGGTCCAGCTGGGCAGAGTGG 843  
 Qy 252 SerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsn 271  
 Db 844 TCACCTGTCTCAACCGCTGTGCGGAGGTGGCAGAGCGCACCGGACCTGCACCAAC 903  
 Qy 272 ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal 291  
 Db 904 CCGCTGCCACTCAACGGAGGGGCTTCTGCGAGGGCGAG---GCATTCCAGAGACCGCC 960  
 Qy 292 SerSerLeuLeuValSerValAspGlySerTrpSerProTrpSerLysLysTrpSerAlaCys 311  
 Db 961 TGCACCACTATGCCAGTCCGATGGGGCGGTGGACGAGTGGAGAGTGGTCCAGCTGC 1020  
 Qy 312 GlyLeuAspCysThrHisTrpArgSerArgGlyCysSerAspProAlaProArgAsnGly 331  
 Db 1021 AGCACTGAGTGTGCCACTGGCGTAGCGCGAGTGCATGGCGCCCCCACCCAGACGGA 1080  
 Qy 332 GlyGluGlyCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal 351  
 Db 1081 GGCGGTGAGTGCAGCGGAGCGTCTCGACTTAAGAACTGCACAGATGGGCTGTGCATG 1140  
 Qy 352 HisSer-----AlaSerGlyProGluAspValAlaLeuTyr 363  
 Db 1141 CAAAGTGAGCTGTCCCGCAGTCTGGAGGCTCAGGG-----GATGGCGGCTGTAT 1194  
 Qy 364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuLeu 382  
 Db 1195 GCGGGGCTCGTGTGGCCATCTTCGTGGTCTGCGCAATCTCATGCGGTGGGGTGGTG 1254  
 Qy 383 ValTyrCysArgLysGlyGluLeuAspSerAspValAlaAspSerSer---IleLeu 401  
 Db 1255 GTGTACCGCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGTCGCCCTG 1314  
 Qy 402 ThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu 421  
 Db 1315 ACTGTGGTTCACCCCGTCACTTTAAGACGCGAAGGCCAGTAAACCGCAGCTCTTA 1374  
 Qy 422 -----ThrIleGlnProAspLeuSerThrThrThr-----ThrTyrGlnGlySerLeuCys 438  
 Db 1375 CACCCCTCTGTGCTCTGACTCTGACCTGACAGCGCGCGGACATCTACCGCGGACCGTGTAT 1434  
 Qy 439 ProArgGluAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerPro 458  
 Db 1435 GCCTTGAGGACTTCC---ACCGCAAAATCCCATGACCAACTCTCTCTGTCTGGACCCC 1491  
 Qy 459 LeuGlyGlyGlyArgHisThrLeuHisHisSerSerProThr----- 472  
 Db 1492 TTACCCAGCCTTAAGTCAAGGCTACAGCTCCAGCACCCAGCGGCTCTGGCGAGGCGCTG 1551  
 Qy 473 SerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyr----- 487  
 Db 1552 GCAGATGGGGTGTACCTGCTGGGGGTCTTGGCGGCTGGGCACATACCTAGCGATTGGCC 1611

Qy 488 -----PheArgSer----- 490  
 Db 1612 CCGGACACCCACTCTCTGCACCTGGCAGCGCAGCCTCGGTTCGCCAGCAGCTCTTGGGCG 1671  
 Qy 491 LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeu 510  
 Db 1672 CTGCCCCGAGACCCAGGAGCAGCTCAGCGGACCTTTGGCTGCTGGTGGGAGGCTC 1731  
 Qy 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530  
 Db 1732 AGCATCCCCCGCACAGTGTGCTGCTGGTGGCCATATGAGCCATTCGCCAGGCAAG 1791  
 Qy 531 IleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeuProLeuAla--- 549  
 Db 1792 TTCTACGAGATGTATCTACTCATCAACAGGAGGAGAAATACC---GTGCCGCTTTCCAGAA 1848  
 Qy 550 GlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyValLeuLeuThr 569  
 Db 1849 GGGACCCACACAGTATTGAGCCCTCGGTGACCTGTGGACCCACAGGCTCTCTGTGTGC 1908  
 Qy 570 ArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerThrSerLeu 589  
 Db 1909 CGCCCCGTATCTCATCCATGCCCATCTGTGCCAAGTCAGTCCCGCTGACTGGATCTTT 1968  
 Qy 590 ArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGlu 609  
 Db 1969 CAGCTCAAGACCCAGGCCCCCAGGAGGCTGGAGCAGAGGTGTGACCTTGGATGAG 2028  
 Qy 610 GluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThr 629  
 Db 2029 GAGACCTCAACACACACCTCTACTGTCAGCTGGAGCCCGGCTGTACATCTCTGCTG 2088  
 Qy 630 GluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArg 649  
 Db 2089 GACCAGCTGGCACCTACGTGTTCAGCGGCGAGTCTATTCCGCTCAGCAGTCAAGCGG 2148  
 Qy 650 LeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGlyTyrAsnIleArgVal 669  
 Db 2149 CTCAGCTCGGCTCTTCGCCCGCCCTCTGCACCTCTCTGGAGTACAGCTTCCGGGTC 2208  
 Qy 670 TyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGlyLysGlnLeu 689  
 Db 2209 TACTGCTGGAGACACGCTGTAGCAGTGAAGAGGTGTGTGAGCTGGAGCGGACTCTG 2268  
 Qy 690 GlyGlyGlnLeuGlnGluProArgValLeuHisPheLysAspSerTyrHisAsnLeu 709  
 Db 2269 GCGGATACTTGTGTGAGGAGCGCAACCGCTAATGTTCAAGGACAGTTACCAACAACCTG 2328  
 Qy 710 ArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeuLeuValSerTyr 729  
 Db 2329 CGCTCTCTCTCATGACCTCCCATGCTCCCATGCTGGAGGAGCAAGCTGTGGCCAAATAC 2388  
 Qy 730 GlnGluLeuProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPhe 749  
 Db 2389 CAGGAGATCCCTTCTATCATATTGGAGTGGCAGCAGAAAGGCCCTCCACTGCATCTTC 2448  
 Qy 750 ThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGln 769  
 Db 2449 ACCCTGGAGGAGCAGCTTGGCTCCACAGAGTCACTTGCAGAGATCTCGTGGCGGCAA 2508  
 Qy 770 ValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPhe 789  
 Db 2509 GTGGAGGGGAGGCGCCAGATATTCAGCTGCATACCATCTGSCA---GAGACACCTGCT 2565  
 Qy 790 AlaGluLeuAlaLeuGluSerGluAlaGly-----ValProAlaLeuValGlyPro 807  
 Db 2566 GGCTCCCTGGACACTCTCTGCTGTGCCCTGGCAGCAGCTGTACACCCAGCTGGGACCT 2625  
 Qy 808 SerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProPro 827  
 Db 2626 TATGCTTCAAGATCCACTGTCCATCCGCGCAGAGATATGCAACAGCTAGATGCCCCC 2685

```

QY      828 CysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeu 847
Db      2686 AACTACGGGGCAATGACTGGCGAGTGTAGCACAGAAGCTCTCTATGACCGGTACCTG 2745
QY      848 SerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArg 867
Db      2746 AATTACTTTGCCACCAAGAGCGAGCCCGGGTGTGATCTGGACCTCTGGGAAGCTCTG 2805
QY      868 HisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnPro 887
Db      2806 CACGACGACATGGGACCTCAACAGCCTGGCGAGTGCCTTGGAGAGATGGGCAAGAGT 2865
QY      888 AspAlaGlyLeuPheThrValSerGluAlaGluCys 899
Db      2866 GAGATGCTGTGGCTGTGGCCACCGAGCGGGAGTGC 2901

```

## RESULT 12

```

US-10-087-684-1
; Sequence 1, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zernhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; PRIOR FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(2857)
US-10-087-684-1
Alignment Scores:

```

```

Pred. No.: 1,72e-251 Length: 2860
Score: 2484.00 Matches: 493
Percent Similarity: 68.30% Conservatives: 149
Best Local Similarity: 52.45% Mismatches: 250
Query Match: 51.89% Indels: 48
DB: 13 Gaps: 18

US-09-970-944-2 (1-899) x US-10-087-684-1 (1-2860)

QY      1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaAlaTrp--- 19
Db      59 ATGGGGGGCCCGAGCGAGCTCGGGCGCGCTGCTGTCGCACTGCTGCTCTGCTGGAC 118
QY      20 -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db      119 CCGAGGCTGAGCAAGACGAGCACTGATTCTGGCAGCGAGGTGCTCCCTGCTTCCCG 178
QY      37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTrpIleVal 56
Db      179 TCAGCGCCAGCAGAGCGCTGCTTCTTCTGAGGAGCCACAGGACGCCCTACATTTG 238
QY      57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db      239 AAGAACAAGCCTGTGGAGCTCCGCTGCCGCGCTTCCCGCCACACAGATCTACTTTCAAG 298
QY      77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db      299 TGCACCGCGAGTGGGTTCAGCCACGACACGTCACACAGGAAGGCCCTGATGAGGCC 358
QY      97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db      359 ACCGCTCTCGGGTGGCGAGGTGGAGTTCAGATCGAGGTGTCGGCAGCAGGAGGAGCTC 418
QY      117 PheGlyLeuGluGluTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys 136
Db      419 TTTGGCTGGAGGATTACTGTGTCAGTGGCTGGCTGGAGTCCGCGCGGACCCACCAAG 478
QY      137 SerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db      479 AGTCGCGAGCCTACGTCCGATCGCTACCTCGCAAGAACTTCGATCAGGAGCCTCTG 538
QY      157 AlaLysGluValSerLeuGluGlnGlnIleValLeuProCysArgProGluGlyIle 176
Db      539 GGCACAGGAGTCCCTTGACCATGAGTTCTCTGAGTGCCTCCGCGCGGAGGAGG 598
QY      177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db      599 CCTGTGGCGAGGTGGAATGGCTCAAGAATGAGATGTCATCGACCCACCCAGGACACC 658
QY      197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db      659 AACTTCTCTGCTCACCATCGACCAACCTCATCTCCGCGAGCGCGCTGTCGACACT 718
QY      217 AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgAtgArgSerAlaSerAlaA 236
Db      719 GCCAACTATACCTGCTGGCCAAAGAACTCTGTGGCCAAACGCCGAGCACCCTGCCACC 778
QY      237 ValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db      779 GTCATCGTCTACGTCAATGGCGGCTGGTCCAGCTGGGCGAGAGTGGTCACTGCTCTCCA 838
QY      257 SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db      839 CGCTGTGGCCGAGGTGGCAAGCGCACCCGAGCTGACCAACCCCGCTCCACTCAAC 898
QY      277 GlyGlyAlaPheCysGluGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db      899 GGAGGGGCTTCTGCGAGGGCCAG---GCATTCCAGAAAGACCGCTGCACCACTATCTGC 955
QY      297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db      956 CCAGTTCGATGGGGCTGGACGAGTGGAGCAAGTGTGTCAGCTGCAGCAGCTGAGTGTGCC 1015

```

```
QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGln 336
Db 1016 CACTGGCGTAGCGCGAGTGCATGGCGCCGCCACCCAGAACGGAGCGGTGACTGCAGC 1075
QY 337 GlyTrpAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db 1076 GGGACGCTGCTGACTCTPAGAACTGCACAGATGGGCTGTGATGATGCACTGGAGGCTCA 1135
QY 356 GlyProGluAspValAlaLeuTyValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db 1136 GGG-----GATCGCGCGTGTATGCGGGGCTCGTGTGGCCATCTTCGTGTGTCGGCA 1189
QY 375 LeuLeuLeuLeuValLeuLeuValTyCysArgGlyLysGluGlyLeuAspSerAsp 394
Db 1190 ATCCCTCATCGCGTGGGGTGGTGTATCCGCGCAACTGCGCTGACTTCGACACAGAC 1249
QY 395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db 1250 ATCACTGACTCATCTGCTGCCCTGACTGTGGTTTCCACCCGCTCAACTTTAAGACGGCA 1309
QY 414 LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr 431
Db 1310 AGGCCAGTAACCGCAGCTCTTACACCCCTCTGTGCTCTCTGACCTGACGACGCGC 1369
QY 432 ---ThrTyrcGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db 1370 GGCATCTACCGCGGACCGGTGTATGCGCTGCAGGACTCC---ACCGACAAATCCCGATG 1426
QY 451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrHisHisSerSer 470
Db 1427 ACCAACTCTCTGTGGACCCCTTACCAGCCTTAAGTCAAGTCTACAGCTCCAGC 1486
QY 471 ProThr-----SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db 1487 ACCACGGGCTCTGGCGAGGCTGGCAGATGGGGCTCACTGTGGGGTCTTTCGCGCCT 1546
QY 485 GlnAsnTy-----PheArgSer----- 490
Db 1547 GGCACATACCTTAGCGATTTCGCGCGGACACCCACTTCTGCACCTGCGCAGCGCCAGC 1606
QY 491 -----LeuProArgGlyThrSerAsnMetThrTyrglyThr 502
Db 1607 CTGGTTCACGAGCTCTTGGGCTGCCCCGAGACCCAGGAGCAGCGGTCAGCGGCACC 1666
QY 503 PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522
Db 1667 TTTGGCTGCTGGGTGGAGGCTCAGCATCCCGGCACAGCGGTCAGCTTGTGTGGGCC 1726
QY 523 ProAspAlaIleProArgGlyLysIleTyrcGlnLeuLeuTyrcGlnLeuHisLysProGlu 542
Db 1727 AATGGAGCCATCCCGAGGCAAGTTCTACGAGATGTAICTCTCATCAACAAAGGCGAA 1786
QY 543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561
Db 1787 AGTACC---CTGCGCGCTTCAGAAAGGACCCACAGACAGTATTGAGCCCTCGGTGACTGT 1843
QY 562 GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581
Db 1844 GGACCCACAGGCTCTGCTGTGCGCGCCGCTCATCTCTCAACATGCCCTGCTGCGGAA 1903
QY 582 ProSerProAspSerTrpSerLeuArgLeuLysLysSerCysGluGlySerTrpGlu 601
Db 1904 GTACGTGCGCGTACTGTGATCTTTACGTCAAGACCCAGGCCACCGAGGCGCTGGAG 1963
QY 602 GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrcGlnGlnLeuGlu 621
Db 1964 ---GAGGTGTGACCTCGATGAGGAGACCCCTGAACACACCTGCTACTGCCAGCTGGAG 2020
QY 622 AlaSerAlaCysTyrcValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641
Db 2021 CCCAGGCGGCTGCATACATCTGCTGGACCGAGCTGCGGACCTTACGTGTTCACGGCGAGTCC 2080
QY 642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661
```

## RESULT 13

```
US-10-087-684-3
; Sequence 3, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
```

```

; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 3
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)...(2857)
US-10-087-684-3

Alignment Scores:
Pred. No.: 1,72e-251 Length: 2860
Score: 2484.00 Matches: 493
Percent Similarity: 68.30% Conservative: 149
Best Local Similarity: 52.45% Mismatches: 250
Query Match: 51.83% Indels: 48
DB: 13 Gaps: 18

US-09-970-944-2 (1-899) x US-10-087-684-3 (1-2860)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaLaTrp--- 19
Db 59 ATGGGGCCCGAGCGGAGCTCGGGCGCGCTGCTGCTGCGACTGCTGCTGCTGCGGAC 118
QY 20 -----LeuArgGlySerGlyAlaGlnGln-----SerAlaThrValAlaAsnProValPro 36
Db 119 CCGAGGCTGAGCAAGCAGGCACTGATCTGCGACGCGAGGTGCTCCTCGTACTCTCCG 178
QY 37 GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrlleVal 56
Db 179 TCAGCGCCAGCAGACGCGCTGCTTCTGTCAGGAGCCACAGGAGCGCTTACATTGTG 238
QY 57 LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db 239 AAGAACAGACCTGTGGAGCTTGGCTGCGCGCGCTTCCCGCCACACAGATCTACTTCAAG 298
QY 77 CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db 299 TGCACGCGGAGTGGTTCAGCCAGAACGACGACGTCACAGAAAGGCGTGGATGAGGCC 358
QY 97 SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnValGlnLysVal 116
Db 359 ACCGCGCTCGGGTGGCGAGGTGCAGATCGAGGTGTCGCGCAGCAGGTGGAGGAGCTC 418
QY 117 PheGlyLeuGluGluTyrlleTrpCysGlnCysValAlaLaTrpSerSerGlyThrThrLys 136
Db 419 TTGGGCTGGAGGATTACTGGTGGCAGTGGCTGGAGCTCGGCTGGAGTCCGCGAGCACC 478
QY 137 SerGlnLysAlaTyrlleArgIleAlaArgLeuArgLysAsnPheGluGlnProLeu 156

```

```

Db 479 AGTGGCCGAGCCTACGTCCGATCGCTACCTGCGCAAGAACTTCCTCAGAGACCTCTG 538
QY 157 AlaLysGluValSerLeuGluGlnIleValLeuProCysArgProGluGlyIle 176
Db 539 GGCAAGGAGGTGCCCTGGACCATGAGGTTCCTCTGAGTGCCTGCCGCGGAGGGGNG 598
QY 177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196
Db 599 CCTGTGGCGAGGTGGAATGGCTCAAGATGAGGATGTCTCATGCCACCCACCCAGGACAC 658
QY 197 AsnValTyrlleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216
Db 659 AACTTCTCTGCTCACCATCGACCAACCTCATCTCCGCGAGCGCTGCTGGGACACT 718
QY 217 AlaAsnTyrlleThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAla 236
Db 719 GCCAACTATACCTGCTGGCCAAAGAACATCGTGGCCAAACGCGGAGCACCCTGCCACC 778
QY 237 ValIleValTyrlleValAsnGlyTyrlleTrpSerThrTrpThrGluTrpSerValCysSerAla 256
Db 779 GTCATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGCAGAGTGTTCACCTGTCTCAAC 838
QY 257 SerCysGlyArgGlyTyrlleGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276
Db 839 CGCTGTGGCCGAGGCTGGCAGAGCGCACCCGAGCTGCACCAACCCCGCTCCACTCAAC 898
QY 277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal 296
Db 899 GGAGGGGCTTCTCGAGGGGCCAG--GCATTCCAGAAAGACCGCTGCACCACTCTGC 955
QY 297 SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr 316
Db 956 CCAGTCGATGGGGCTGGAGGAGTGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCC 1015
QY 317 HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336
Db 1016 CACTGGCGTAGCGGAGTGATGGCGGCCCCACCCAGAACGAGGCGCTGATGTCAGC 1075
QY 337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355
Db 1076 GGGAGCGTGTCTGACTCTAAGAACTGCACAGATGGGTGTGTCATCAACTGGAGGCTCA 1135
QY 356 GlyProGluAspValAlaLeuTyrlleValGlyLeu---IleAlaValAlaValCysLeuVal 374
Db 1136 GGG-----GATGGCGCGCTGTATCGGGGCTCGTGGTGGCCATCTTGTGTCGTGCA 1189
QY 375 LeuLeuLeuLeuValIleLeuValTyrlleValTyrlleValLysGluGlyLeuAspSerAsp 394
Db 1190 ATCCTCATGGCGGTGGGGTGGTGTGTACCGCCCAACTGCCGTGACTTCGACACGAC 1249
QY 395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413
Db 1250 ATCACTGACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1309
QY 414 LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThrThr 431
Db 1310 AGGCCAGTAACCCGACGCTCTACACCCCTCTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1369
QY 432 ---ThrTyrlleGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450
Db 1370 GGCATCTACCGGAGCGCTGTATGCTCGCCTGCGAGCTCC---ACGACAAATATCCCAATG 1426
QY 451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeuHisSerSer 470
Db 1427 ACCAACTCTCTCTGCTGAGCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGCTCCAGC 1486
QY 471 ProThr-----SerGluAlaGluGluPheValSerArgLeuSerThr 484
Db 1487 ACCAGGGCTCTGGGCCAGGCTGCGAGATGGGGTGCAGTGGGGTCTTGGGGCTTGGCGCT 1546
QY 485 GlnAsnTyrlle-----PheArgSer----- 490

```

Db	1547	GCACATACCTAGCGATTTCGCCGGGACACCCACTTCTGCACCTGCGCAGCGCCAGC	1606
QY	491	-----LeuProArgGlyThrSerAsnMetThrTyrGlyThr	502
Db	1607	CTCGGTCCACGAGCTCTTGGCCCTGCGCCGAGACCCAGGAGCAGCGCTCAGCGGCACC	1666
QY	503	PheAsnPheLeuGlyClyArgLeuMetIleProAsnThrGlyIleSerLeuIlePro	522
Db	1667	TTTGCTGCTGCTGGTGGGAGCTCAGCATCCCGGCACAGGGGTGCTGTGTGTCGCC	1726
QY	523	ProAspAlaIleProArgGlyIleTyrGluIleTyrLeuThrLeuHisIleValProGlu	542
Db	1727	ANTGAGGCAATCCCGAGGCAAGTCTACGAGATGTATCTACTCATCAACAGGCAGAA	1786
QY	543	AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys	561
Db	1787	AGTACC---CTGCGCTTTTCAGAAAGGACCCAGACAGATATTAGCCCTCGGTGACCTGT	1843
QY	562	GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu	581
Db	1844	GGACCCACAGGCTCTCTGTGTGCCGCCCTCATCTCACCATGCCCTCGCTGCGCAA	1903
QY	582	ProSerProAspSerTrpSerLeuArgLeuLysGlnSerCysGluGlySerTrpGlu	601
Db	1904	GTACGTGCCCTGACTGGATCTTTACGCTCAAGACCCAGGCCACCAGGGCCACTGGAG	1963
QY	602	GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu	621
Db	1964	--GAGTGTGACCTCGATGATGAGGAGACCTCAACACACACCTGTACTGCCAGCTGGAG	2020
QY	622	AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla	641
Db	2021	CCAGGGCCCTGTCACATCTGCTGGACGACTGGGACCTTACGTGTTCAGGGCGAGTCC	2080
QY	642	LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThr	661
Db	2081	TATTCGGCTCAGCAGTCAAGCGCTCCAGCTGGCGCTCTTGCCCCCGGCCCTCTGCACC	2140
QY	662	SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu	681
Db	2141	TCGCTGGAGTACAGCTCCGGTCTACTGCTGGAGCACAGCGCTGTAGCACTGAAGGAG	2200
QY	682	ValValGlnLeuGlyLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHis	701
Db	2201	GTCTGGAGTGGAGCGGACTTGGGGGATCTTGTGGAGGAGCGAAACCGCTAATG	2260
QY	702	PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp	721
Db	2261	TTCAAGGACAGTTTACCACAACTCGCGCTCTCCCTCCATCAGCTCCGCCATGCCATTGG	2320
QY	722	LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThr	741
Db	2321	AGGAGCAAGCTGCTGGCCAAATACAGGAGATCCCTTCTATCACATTTGAGTGGCAGC	2380
QY	742	GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu	761
Db	2381	CAGAGGCCCTCCACTGACTTTACCCCTGGAGAGGCAGAGCTGGCTCCACAGAGCTC	2440
QY	762	AlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe	781
Db	2441	ACCTGCAAGATCTCGTGGCGCAAGTGGAAAGGGAGGGCCAGATATTCAGCTGCATACC	2500
QY	782	AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuSerGluAlaGly---	800
Db	2501	ACTGTGCA---GAGACACCTGTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGC	2557
QY	801	---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys	819
Db	2558	ACTGTACCCACCCAGCTGGGACCTTATGCTCTCAAGATCCCACTGTCTCCATCCGCCAGAAG	2617
QY	820	IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln	839
Db	2618	ATATGCAACAGCTTAGATGCCGCCCACTACGGGGCAATGACTGGCGGATGTTAGCACAG	2677
QY	840	LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet	859
Db	2678	AAAGCTCTCTATGACCGGTACTTGAATTATTTGCCCAAGCGAGCCCGGCTGTG	2737
QY	860	IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla	879
Db	2738	ATTCCTGGACCTCTGGGAAGCTCTGCAGCAGGAGCATGGGAGCTCAACAGCCTGGCGAGT	2797
QY	880	AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys	899
Db	2798	GCCTTGAGGAGATGGCAAGAGTGAGATGCTGTGGTGTGGCTGCTGCGCACCGAGCGGGACTGC	2857
RESULT 14			
US-10-218-779-1			
; Sequence 1, Application US/10218779			
; Publication NO. US20040029222A1			
; GENERAL INFORMATION:			
; APPLICANT: Edinger, Shlomit			
; APPLICANT: MacDougall, John			
; APPLICANT: Millet, Isabelle			
; APPLICANT: Ellerman, Karen			
; APPLICANT: Stone, David			
; APPLICANT: Gerlach, Valerie			
; APPLICANT: Alsobrook II, John			
; APPLICANT: Lepley, Denise			
; APPLICANT: Rieger, Daniel			
; APPLICANT: Burgess, Catherine			
; APPLICANT: Casman, Stacie			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Boldog, Ferenc			
; APPLICANT: Li, Li			
; APPLICANT: Fadigar, Muralidhara			
; APPLICANT: Mishra, Vishnu			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Shenoy, Suresh			
; APPLICANT: Rastelli, Luca			
; APPLICANT: Tchernev, Velizar			
; APPLICANT: Vernet, Corine			
; APPLICANT: Zerhusen, Bryan			
; APPLICANT: Malyankar, Uriel			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Miller, Charles			
; APPLICANT: Gangolli, Saha			
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
; FILE REFERENCE: 21402-214			
; CURRENT APPLICATION NUMBER: US/10/218,779			
; CURRENT FILING DATE: 2002-08-14			
; PRIOR APPLICATION NUMBER: 60/253,834			
; PRIOR FILING DATE: 2000-11-29			
; PRIOR APPLICATION NUMBER: 60/250,-926			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: 60/264,180			
; PRIOR FILING DATE: 2001-01-25			
; PRIOR APPLICATION NUMBER: 60/313,656			
; PRIOR FILING DATE: 2001-08-20			
; PRIOR APPLICATION NUMBER: 60/327,456			
; PRIOR FILING DATE: 2001-10-05			
; NUMBER OF SEQ ID NOS: 216			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 2860			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-218-779-1			
Alignment Scores:			
Pred. No.:		1-72e-251	Length: 2860
Score:		2484.00	Matches: 493
Percent Similarity:		68.30%	Conservative: 149
Best Local Similarity:		52.45%	Mismatches: 250
Query Match:		51.89%	Indels: 48



DB:	13	Gaps:	18
US-09-970-944-2 (1-899) x US-10-218-779-1 (1-2860)			
QY	1	MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyIleValLeuAlaIaTrp---	19
DB	59	ATGGGGCGCGGAGCGGAGCTCGGGCGCGCTGCTGCTGCACTGCTGCTGCTGGAC	118
QY	20	-----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAlaAsnProValPro	36
DB	119	CGAGGCTGAGCCAGCAGCAGCTGATCTGGCAGCGAGTGCTCCTGACCTCTCCCG	178
QY	37	GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValIleVal	56
DB	179	TCAGCGCCAGCAGAGCGCTGCGCTACTTCTTCAGAGCCACAGACGCCCTACATTGTG	238
QY	57	LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys	76
DB	239	AGAACAAGCGCTGTGAGCTCGGCTCGCGCGCTTCCCGCCACACAGATCTACTCAAG	298
QY	77	CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer	96
DB	299	TGCACCGCGAGTGGTTCAGCAGAACGACACGTCACACAGGAAGCGCTGGATGAGGCC	358
QY	97	SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal	116
DB	359	ACCGGTCTCGGCTGGCGAGTGCAGATCGAGGTCTCGCGCAGCAGGTGGAGGCTC	418
QY	117	PheGlyLeuGluGluTrpTrpCysGlnCysValAlaTrpSerSerSerGlyThrThrLys	136
DB	419	TTTGGCTGGAGATTACTTGGTCCAGTGGCTGCGCTGGAGTTCGCGGCGCACCAAG	478
QY	137	SerGlnLysAlaTrpIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu	156
DB	479	AGTGGCGGAGCTTACCTCGCATGCTACCTCGCGCAAGACTTCGATCAGAGGCTCTG	538
QY	157	AlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProGluGlyIle	176
DB	539	GGCAAGGAGTGGCCCTGGACCATGAGTTCTCTGCACTGCGCGCGCGGAGGGGGT	598
QY	177	ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro	196
DB	599	CCTGTGGCGAGTGGATGGTCAAGATGAGGATGCTCATCGACCCACCCAGGACACC	658
QY	197	AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr	216
DB	659	AACTTCCTGCTCACCATCGACACACACCTCATCCCGCAGCGCGCTGCGGACACT	718
QY	217	AlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAlaAla	236
DB	719	GCCAACTATACCTGGTGGCCAAAGAACATCGTGGCCAAACGCCGGAGCACCTGCCACC	778
QY	237	ValIleValTyrValAsnGlyTrpSerThrTrpThrGluTrpSerValCysSerAla	256
DB	779	GTCACTCGTCTAGCTGAATGGCGGCTGGTCCAGCTGGCGAGAGTGGTCACTGCTCCAAC	838
QY	257	SerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn	276
DB	839	CGCTGTGGCGGAGGCTGGCAGAGCGCACCCCGAGCTGCACCAACCCGCTCCACTCAAC	898
QY	277	GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuVal	296
DB	899	GGAGGGGCTTCTCGAGGGCCAG---GCATTCCAGAGACCGCTGCACCACTCTGC	955
QY	297	SerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThr	316
DB	956	CCAGTCGATGGGCGTGGACGAGTGGAGCAAGTGGTCACTGCGAGCACTGAGTGGCC	1015
QY	317	HisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln	336
DB	1016	CAGTGGCTAGCGGAGTGCATGGCGGCCACCCACCCAGAGCGGCGCTGACTGCAGC	1075
QY	337	GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer	355
DB	1076	GGGACGCTCTGACTCTAAGAACTCCACAGATGGCTGTGATGCACTGGAGGGCTCA	1135
QY	356	GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal	374
DB	1136	CGG-----GATGGCGCGCTGATGCGGGCTCGTGGTGGCCATCTTCGTGGTGGCA	1189
QY	375	LeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeuAspSerAsp	394
DB	1190	ATCCCTATCGCGGTGGGGTGGTGTATCCGCCCACTGCGCGTGACTTCGACACAGAC	1249
QY	395	ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer	413
DB	1250	ATCACTGACTCATCTGCTCCTGACTGTGTGTTCACCCCGTCACTTTAAGACGGCA	1309
QY	414	LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThrThr	431
DB	1310	AGGCCAGTAACCCGCGAGCTCTACACCCCTCTGTGCTCTCTGACCTGACAGCGCGCC	1369
QY	432	---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu	450
DB	1370	GGCATCTACCGCGGACCGGTGATGCCCTGCGAGGACTCC---ACCGACAAATCCCCATG	1426
QY	451	ThrAsnGlyHisLeuLeuSerProLeuGlyGlyGlyArgHisThrLeuHisHisSerSer	470
DB	1427	ACCACTCTCTCTGCTGGACCCCTTACCAGCCCTTAAGGTCAAGTCTACAGTCCAGC	1486
QY	471	ProThr-----SerGluAlaGluPheValSerArgLeuSerThr	484
DB	1487	ACCGGGCTCTGGCGCAGGCTGGCAGATGGGCTGACCTGCTGGGGTCTTTGGCGCT	1546
QY	485	GlnAsnTyr-----PheArgSer-----	490
DB	1547	GGCACAATACCTAGCGATTTCGCCGGGACACCCACTTCTCGACCTGCGCAGCGCCAGC	1606
QY	491	-----LeuProArgGlyThrSerAsnMetThrTyrGlyThr	502
DB	1607	CTCGGTTCCACGACCTCTTGGGCTGCCCCGAGACCCAGGAGGAGCGTACGCGGCACC	1666
QY	503	PheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuLeuPro	522
DB	1667	TTTGGCTGCTGGTGGGAGGCTCAGCATCCCGGCACACGGGTGCTGCTGGTGGCC	1726
QY	523	ProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGlu	542
DB	1727	AATGAGCGCATTCGCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAGGACGAA	1786
QY	543	AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys	561
DB	1787	AGTACC---CTGCCGCTTCAGAGGGACCCAGACAGATATTGAGCCCTCGGTGACCTGT	1843
QY	562	GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu	581
DB	1844	GGACCCACAGGCTCCTGCTGGCGCGCGCTCATCTCACCATCCCCACCTGTGCCGAA	1903
QY	582	ProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGlySerTrpGlu	601
DB	1904	GTCACTGGCGGTGACTGGATCTTTTCAGCTCAAGACCCAGGCGCCACGAGGCGCTGGAG	1963
QY	602	GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu	621
DB	1964	---GAGGTGGTGACCTCGATGAGGAGACCCCTGAACACACACCTGCTACTGCCAGTGGAG	2020
QY	622	AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla	641
DB	2021	CCAGGGCGCTGCACATCTGCTGGACCGAGCTGGGCACCTACGTGTTACGSGCGAGTCC	2080
QY	642	LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThr	661
DB	2081	TATTCCCTCAGCAGTCAAGCGGTCCAGTGGCGCTTTCGCCCGCGCTCTGCTGCACC	2140
QY	662	SerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu	681

```

Db      2141  TCCTGGAGTAGAGCTCGGGTCTACTGCTGGAGGACACGCCCTGTAGACTGAGGAG 2200
QY      ValValGlnLeuGluysGlnLeuGlyGlnLeuLeuLeuLeuLeuLeuLeuHis 701
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2201  GTGCTGGAGCTGGAGCGACTCTGGCGGATACTTGGTGGAGGAGCGAAACCGCTAATG 2260
QY      PheLysAspSerTyHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrp 721
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2261  TTCAAGGAGAGTTACACAACTGGCCCTCTCCCTCATGACCTCCCATGCCAATGG 2320
QY      LysSerLysLeuLeuValSerTyHisGlnGluLeuPheProPheTyHisIlePheAsnGlyThr 741
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2321  AGGAGCAAGCTGCTGCCAAATACAGAGATCCCTTTATACATTTGGATGGCAGC 2380
QY      GlnArgTyLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2381  CAGAAGGCCCTCCACTGCATCTTACCTGGAGGACAGACTTGGCCTCCACAGAGCTC 2440
QY      AlaCysLeuLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2441  ACTGCAAGATCTGCTGGCGCAAGTGAAGGGAGGGCCAGATATTCAGCTGCATACC 2500
QY      AsnIleThrLysAspThrArgPheAlaGluLeuAlaLeuLeuSerGluAlaGly--- 800
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2501  ACTCTGGCA--GAGACACCTGCTGGCTCCCTGGACACATCTCTGCTCTGCCCTGGCAGC 2557
QY      ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuLeuArgGlnLys 819
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2558  ACTGTCAACACCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAAG 2617
QY      IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGln 839
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2618  ATATGCAACAGCTAGATGCCCAACTACAGGGGCAATGACTGGCGGATGTAGCACAG 2677
QY      LysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMet 859
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2678  AGCTCTCTATGACCGGTACCTGAATTAATTGTCACCAAGCGAGGCCACGGGTG 2737
QY      IleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAla 879
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2738  ATCTGGACTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCCTGGCGAGT 2797
QY      AlaValAlaGlyLeuGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      2798  GCCTTGGAGAGATGGGCAAGAGTAGATGCTGCTGGCTGGCCACCGGAGGAGTGC 2857

RESULT 15
US-10-218-779-3
; Sequence 3, Application US/10218779
; Publication No. US2004002922A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Sytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shenoy, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine

```

```

; APPLICANT: Zerhusen, Bryan
; APPLICANT: Malyankar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangolli, Esha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,-926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-218-779-3

Alignment Scores:
Pred. No.: 1,72e-251      Length: 2860
Score: 2484.00           Matches: 493
Percent Similarity: 68.30%      Conservative: 149
Best Local Similarity: 52.45%    Mismatches: 250
Query Match: 51.89%            Indels: 48
DB: 13                      Gaps: 18

US-09-970-944-2 (1-899) x US-10-218-779-3 (1-2860)
QY      1  MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeuValLeuAlaAlaTrp--- 19
Db      59  ATGGGGCGCGGAGCGAGCTCGGGCGCGCTGCTGCTGGCAGCTGCTCTCTCTCTGGGAC 118
QY      20  -----LeuArgGlySerGlyAlaGlnGln---SerAlaThrValAlaAsnProValPro 36
Db      119  CCGAGGCTGAGCAACAGCAGGCACTGATTCTGGCAGCGAGGTGCTCCTGACTCTCTCCCG 178
QY      37  GlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspValTyrlleVal 56
Db      179  TCAGCGCCACGACGAGCGCTGCCCTACTTCTCGAGGAGCCACAGGACGCTTACATTGTG 238
QY      57  LysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLys 76
Db      239  AAGAACAAGCTGTGGAGCTTCGTCGCGGCCCTCCCGCCACACAGATCTACTTCAAG 238
QY      77  CysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThrAspGlySer 96
Db      299  TGCAACGCGGAGTGGGTGAGCCAGCAACAGCAGCTCACAGGAGGCGCTGGATGAGGCC 358
QY      97  SerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysVal 116
Db      359  ACCGGCTCGGGGTGCGAGGTGCGAGTGCAGATCGCGGAGGAGTGGAGGAGCTC 418
QY      117  PheGlyLeuGluGluTyrlTrpCysGlnCysValAlaTrpSerSerGlyThrThrLys 136
Db      419  TTTGGCTGGAGATTACTGGTGCAGTGGCTGGCTCGGAGCTCCGAGGACCCACCAAG 478
QY      137  SerGlnLysAlaTyrlleArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeu 156
Db      479  AGTCGCGGAGCTTACGTCGCGCATCGCTACCTCGCAAGAACTTCGATCAGGAGCCTCTG 538
QY      157  AlaLysGluValSerLeuGlnGlnGlyIleValLeuProCysArgProGluGlyIle 176
Db      539  GGCAAGGAGGTGCCCTGGACCATGAGGTTCCTCTGAGTGCCTCCCGCGGAGGGGTG 598

```

177 ProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspPro 196  
Db CTTGTGGCGAGGTGAATGGCTCAAGATGAGATGTCATGACCCACCGAGACACC 658  
197 AsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThr 216  
Db AACCTTCCTGCTACCATCGACCAACAACCTCATCGCCGAGCGCGCTGTGGACACT 718  
217 AlaAsnTyrThrCysValAlaIleAsnIleValAlaArgArgSerAlaSerAlaAla 236  
Db GCCAATAATACCTGCGTGGCCAGAAACATCGTGGCCAAAGCCGCGAGCACCATGCGCAC 778  
237 ValIleValTyrValAsnGlyGlyTyrSerThrTyrThrGluTyrSerValCysSerAla 256  
Db GTCATCGTCTACGTGAATGGCGCTGTCACAGCTGGGCAGAGTGTCAACCTGCTCCAAC 838  
257 SerCysGlyArgGlyTyrGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsn 276  
Db CGCTGTGGCGGAGGCTGGCAGAAAGCCACCCGACCTGCACCAACCCCGCTCCACTCAAC 898  
277 GlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuVal 296  
Db GAGGGGCGCTTTCGAGGCGCAG----GCATTCAGAAAGACCGCCTGCACCATCTGC 955  
297 SerValAspGlySerTyrProTyrSerTyrSerLysTyrSerAlaCysGlyLeuAspCysThr 316  
Db CCAGTGCATGGGCGTGGACGAGTGGAGCAAGTGTTCAGCCTGCAGCAGCTGAGTGTGCC 1015  
317 HisTyrArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluCysGln 336  
Db CACTGCGGTAGCGCGAGTGCATGGCGGCCGCCACCCAGAACCGAGCGCGTGACTGCAGC 1075  
337 GlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHis---SerAlaSer 355  
Db GGGACCTGCTCGACTCTAGAACTCCACAGATGGGCTGTGTGATCAACTGGAGGCGCTCA 1135  
356 GlyProGluAspValAlaLeuTyrValGlyLeu---IleAlaValAlaValCysLeuVal 374  
Db GGG-----GATGCGGCGGTGATGCGGGCTCGTGGTGCCATCTTCGTGTGGTGGCA 1189  
375 LeuLeuLeuLeuValLeuIleLeuValTyrCysArgLysGluGlyLeuAspSerAsp 394  
Db ATTCCTCATGGCGGTGGGGTGGTGTGTACCGCGCAACTGCGCTGACTTCGACACAGAC 1249  
395 ValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSerIleLysProSer 413  
Db ATCACTGATCTATCTGCTCCCTGACTGTGTGTTCACCCCGTCAACTTAAGACGGCA 1309  
414 LysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThrThr 431  
Db AGGCCAGTAACCGCGAGCTCTACACCCCTCTGTGCTCTCTGACTGACAGCCAGCGCC 1369  
432 ---ThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGlnLeu 450  
Db GGCATCTACCGCGGACCCGTGTATGCTGCTGCAGGACTCC---ACCGACAAAATCCCATG 1426  
451 ThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeuHisSerSer 470  
Db ACCAACTCTCTGCTGAGCCCTTACCACCCCTTAAGGTCAAGTCTACAGCTCCAGC 1486  
471 ProThr-----SerGluAlaGluPheValSerArgLeuSerThr 484  
Db ACCACGGGCTCTGGGCCAGGCTGGCAGATGGGCTGACCTGCTGGGGTCTTGGCGCT 1546  
485 GlnAsnTyr-----PheArgSer----- 490  
1547 GGCACATACCTAGCGATTTCGCCCGGACACCCCACTTCCTGCACCTGGCGAGCGCCAGC 1606  
491 -----LeuProArgGlyThrSerAsnMetThrTyrGlyThr 502  
1607 CTCGGTCCACAGAGCTCTTGGGCGCTGCCCCAGAGCCAGGAGGAGCGGCTCAGCGCAC 1666  
503 PheAsnPheLeuGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIlePro 522

1667 TTTGGCTGCTGGTGGGAGGCTCAGCATCCCGGCACAGGGGTGAGTTCGTGTGGTCCC 1726  
523 ProAspAlaIleProArgGlyLysIleTyrGluIleTyrIleThrLeuHisLysProGlu 542  
Db AATGGAGCCATTCGCCAGGCGAAGTCTTACGAGATGTATCTACTCATCAACAGGCAGAA 1786  
543 AspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSerProIleValSerCys 561  
Db AGTACC---CTGCCGCTTTCAGAAAGGAGCCACAGACAGTATTGAGCCCTCGTGGTGT 1843  
562 GlyProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGlu 581  
Db GGAACCCACAGGCTCTGTGTGCCGCCCGCTCATCTCACCATGCCCATCTGTGCCGAA 1903  
582 ProSerProAspSerTyrSerLeuArgLeuLysLysGlnSerCysGluGlySerTyrGlu 601  
Db GTCAGTGGCGGTGACTGGATCTTTCAGCTCAAGACCCAGGCCACCCAGGCGCACCTGGAG 1963  
602 GlnAspValLeuHisLeuGlyGluAlaProSerHisLeuTyrTyrCysGlnLeuGlu 621  
Db ---GAGGTGGTACCCCTGGATGAGGAGACCTTGAACACACCCCTGTCTACTGCCAGCTGGAG 2020  
622 AlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAla 641  
Db CCCAGGCGCTGTACATCTCTGTGCACGCTGGGCAGCTAGCTGTTCACGGCGAGTCC 2080  
642 LeuSerValAlaAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThr 661  
Db TATTCGCTCAGCAGTCAAGCGGTCCAGCTGGCGGTCTTCGCCCGCCCTCTGCACC 2140  
662 SerLeuGlyTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGlu 681  
Db TCCCTGGAGTACAGCTCCGGTCTACTCGCTGGAGGACACGCTGTACACTGAAGGAG 2200  
682 ValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnProArgValLeuHis 701  
Db GTGCTGGAGCTGGAGGGAGCTCTGGCGGATCTTGGTGGAGGAGCCGAAACCGCTAATG 2260  
702 PheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTyr 721  
Db TTCAAGGACAGTTACCACAACCTGCGCTCTCCCTCCATGACCTCCCATGCCATTGG 2320  
722 LysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTyrAsnGlyThr 741  
Db AGGACAAAGCTCTGCGCAATACCAGGAGATCCCTCTCATCATTTTGGAGTGGCAGC 2380  
742 GlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeu 761  
Db CAGAGGCCCTCCATGACCTTTCACCTGGAGAGGCACAGCTTGGCTTCCACAGAGTCC 2440  
762 AlaCysLysLeuTyrValTyrGlnValGluGlyAspGlyGlnSerPheSerIleAsnPhe 781  
Db ACCTCAAGATCTGCGTGGCAAGTGAAGGGGAGGCGCAGATATTCAGCTGCATACC 2500  
782 AsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGly--- 800  
Db ACTCTGGCA---GAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTCGCAGC 2557  
801 ---ValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLys 819  
Db ACTGTACACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGCTCCATCGCCAGAG 2617  
820 IleIleSerSerLeuAspProProCysArgArgGlyAlaAspTyrArgThrLeuAlaGln 839  
Db ATATCAACAGCTAGATGCCCCCAACTCAGGGGCAATGACTGGCGGATGTATGACAG 2677  
840 LysLeuHisLeuAspSerHisLeuSerPheAlaSerLysProSerProThrAlaMet 859  
Db AAGCTCTATGGACCGGTACTGAAATTAATTGTCACCAAGAGCGAGGCCACCGGCTGTG 2737  
860 IleLeuAsnLeuTyrGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAla 879

Db 2738 ATCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGT 2797  
Qy 880 AlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGluCys 899  
|||::: ::|||::: :::  
Db 2798 GCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGC 2857

Search completed: October 6, 2004, 22:57:28  
Job time : 1147 secs



Db 361 CCKRGNKSKPLKPKQNMSEKAGGIYSEPPGVRRLLEHGHQHTLLGKISCSQYFEP 420  
 QY 408 VSIKPSKADNPHTLLTQDPLSTTTTGGSLCPQDGPSPKFQLTNGHLLSPGLGGHRTLLH 467  
 Db 421 PPL-----PHSTTLR---SGKSAPSGYSTFNAGSRAALIQECSSSSSGSGGKRTMLR 470  
 QY 468 HSSPT-SEAEFVSRSLSTQNYFRSLPRGTS-NMTYGTFFNLGLGRMLPNTGISLLIPDDA 525  
 Db 471 TSSSNCDDNNATLYDMEDKSVLGLDTSQNTVAQAQIDNSGARLSLSKSGARLIPELA 530  
 QY 526 IPGKTYEYLTLLHKPEDVRLPLAGCQTLSPVSCGPGV-----LLTRPVILAMDH 578  
 Db 531 VEGEKM--LYLAVSDTLTQPHLKPIESALSPVIVIGQCDVMSAHDNLRFPVVSFRH 588  
 QY 579 CGEPPSP-DSWSRLKQKQCEGSEWQVHLHGEAPSHLYYQCLEASA-----CYVFT 629  
 Db 589 CASTFPDRNQWFTL--YADEGSGWQKAVTIGENLNTNMVQPEQPKKNDGFGWCHVMT 646  
 QY 630 EQLGRLALYGEAL--SVAAAKRLKLLLPAPVACTSLE--YNIRVYCLHDTHDALKKEVQL 695  
 Db 647 YSLARLMLAGHPRRNSLSAAKRVHLAVFGPTEMSAYRRPFELRVYCVPETGAAMESVWKQ 706  
 QY 686 EKOLGGQLIQEPR--VLHFKDSYHNLRLSLHDV-PSSLWKSKLLVSVQELPFVHIWNGTQ 742  
 Db 707 ED--GSRLLCESNDFILNEKG---NLCICIEDVIPGFCSDGPEVHISETOHRFV--AQ 758  
 QY 743 RYLHCTFTTLERSVSPSTDACKLWVQVEGDGQSGFSINFNITKTRPAELALLALESEAGVP 802  
 Db 759 NGLHCSLKFRPKPEINGSQFSTRVIVYQKASSTPEVMW--EVSNEPELYDATSEEREKGSV 816  
 QY 803 ALVGPSPAPKPIFLIRKQIISLDPCCRGADWTLAQKHLDSHLHFFASKP--SPTAMI 860  
 Db 817 CV-----EPLPFGVKDELARLLDMPNESHSDRWGLAKKHYLYLOFFASFPDCSPTSL 872  
 QY 861 LNLWEARHPFNGN-LSOLAAAVAGLQOPDA 889  
 Db 873 LDLWEASSSGSARAVPDLLOTLRVMGEPDA 902

RESULT 2  
 B44294  
 unc-5 protein, long form - Caenorhabditis elegans  
 N;Contains: unc-5 protein, short form  
 C;Species: Caenorhabditis elegans  
 C;Date: 30-Apr-1993 #sequence revision 28-Jul-1995 #text\_change 05-Nov-1999  
 C;Accession: B44294; T32540; A44294  
 R;Leung-Hagetein, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.; Hedgecock, E.M.;  
 Cell 71, 289-299, 1992  
 A;Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 do  
 A;Reference number: A44294; PMID:93046629; PMID:1384987  
 A;Contents: variety Bergerac  
 A;Accession: B44294  
 A;Molecule type: DNA  
 A;Residues: 1-947 <LEU>  
 A;Cross-references: GB:S47168; NID:G258527; PIDN:AAB23867.1; PID:G258529  
 A;Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672,  
 A;Note: authors translated the codon CTA for residue 642 as Val; sequence shown follows  
 A;Note: mRNA lacking the first exon is equally prevalent  
 R;Latrelle, P.  
 submitted to the EMBL Data Library, December 1997  
 A;Description: The sequence of C. elegans cosmid B0273.  
 A;Reference number: Z21187  
 A;Accession: T32540  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-947 <LAT>  
 A;Cross-references: EMBL:AF036698; PIDN:AAB88356.1; GSPDB:GN00022; CESP:B0273.4b  
 A;Experimental source: strain Bristol N2; clone B0273  
 C;Genetics:  
 A;Gene: unc-5  
 A;Map position: 4  
 A;Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3  
 C;Function:

A;Description: required for guidance of pioneering axons and cells migrating dorsally al  
 C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type  
 C;Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane pro  
 F;30-947/Product: unc-5 protein, short form #status predicted <ALT>  
 F;46-116/Domain: immunoglobulin homology <IM1>  
 F;153-211/Domain: immunoglobulin homology <IM2>  
 F;229-300/Domain: thrombospondin type 1 repeat homology #status atypical <THR1>  
 F;301-354/Domain: thrombospondin type 1 repeat homology <THR2>  
 F;365-390/Domain: transmembrane #status predicted <TMW>  
 F;512-559/Domain: SH3 homology <SH3>  
 F;53-114, 65-112, 160-209/Disulfide bonds: #status predicted  
 F;206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 950.5; DB 1; Length 947;  
 Best Local Similarity 28.5%; Pred. No. 1.7e-61;  
 Matches 265; Conservative 165; Mismatches 375; Indels 125; Gaps 31;

QY 49 EPEDVIVNKNPVLVCKAVPATQFFKCNQWVRQVDHVBIE--STDGSSGSGPTMEVRI 106  
 Db 37 QPAGGVIRNKPRLQCRANHATIKYKSSKWD--DSRIEKLIGTSTSGVGYDASV 94  
 QY 107 NVSRQQVEKVFGLLEYWCQCAWSSSG-----TTKSQKAYIRIARLKNFQEPPLAKEVS 161  
 Db 95 DISRIDVDTSGHDAFQCQCYA--SGDDDDQDVASDVATVHLAYWKHFLKSPVQVRVQ 151  
 QY 162 LQGVILPCRPPEGPPIPAEVENLREDLVDPDLDPNVYITREHSLVVRQARLADTANYTC 221  
 Db 152 EGTTLQPCAPESDPKAEITWYKGVVQVP--DANVIRASDGLSMAARLSDSNGYT 209  
 QY 222 VAKNIVARRRSAAVIVVNGWSTWTW-SVCSASCG-----RGWQKR 265  
 Db 210 EATNVANSKTDPEVQIIVDGSNWSWPIGTCHVDCPLLRQHAHRIHPDVLPHQRR 269  
 QY 266 SRSCNTPAPLNGAFCEGQNVHDRTVSSLLSVSDGSPWSKWSACGLDCTHWSRECS 325  
 Db 270 TRTCNNPAPLNDGEYCKGEBEEMTRS--CKVPCKLDGWSWSWDSACSSSCHRYRTRACTV 328  
 QY 326 PAPRNGCEBQCTDTRNCTSDLCVHSAG--PEDVALYVGLIAVAVCLVLLVLLV 383  
 Db 329 PPPMNGQPCFGDMLTQCPAQLCTADSSRIVISDTAVYGVSVASIFVASFILAILAMF 388  
 QY 384 YCR-----KKEGLDSDVADS-----SILTSQFOP 407  
 Db 389 CCKRGNKSKPLKPKQNMSEKAGGIYSEPPGVRRLLEHGHQHTLLGKISCSQYFEP 448  
 QY 408 VSIKPSKADNPHTLLTQDPLSTTTTGGSLCPQDGPSPKFQLTNGHLLSPGLGGHRTLLH 467  
 Db 449 PPL-----PHSTTLR---SGKSAPSGYSTFNAGSRAALIQECSSSSSGSGGKRTMLR 498  
 QY 468 HSSPT-SEAEFVSRSLSTQNYFRSLPRGTS-NMTYGTFFNLGLGRMLPNTGISLLIPDDA 525  
 Db 499 TSSSNCDDNNATLYDMEDKSVLGLDTSQNTVAQAQIDNSGARLSLSKSGARLIPELA 558  
 QY 526 IPGKTYEYLTLLHKPEDVRLPLAGCQTLSPVSCGPGV-----LLTRPVILAMDH 578  
 Db 559 VEGEKM--LYLAVSDTLTQPHLKPIESALSPVIVIGQCDVMSAHDNLRFPVVSFRH 616  
 QY 579 CGEPPSP-DSWSRLKQKQCEGSEWQVHLHGEAPSHLYYQCLEASA-----CYVFT 629  
 Db 617 CASTFPDRNQWFTL--YADEGSGWQKAVTIGENLNTNMVQPEQPKKNDGFGWCHVMT 674  
 QY 630 EQLGRLALYGEAL--SVAAAKRLKLLLPAPVACTSLE--YNIRVYCLHDTHDALKKEVQL 685  
 Db 675 YSLARLMLAGHPRRNSLSAAKRVHLAVFGPTEMSAYRRPFELRVYCVPETGAAMESVWKQ 734  
 QY 686 EKOLGGQLIQEPR--VLHFKDSYHNLRLSLHDV-PSSLWKSKLLVSVQELPFVHIWNGTQ 742  
 Db 735 ED--GSRLLCESNDFILNEKG---NLCICIEDVIPGFCSDGPEVHISETOHRFV--AQ 786  
 QY 743 RYLHCTFTTLERSVSPSTDACKLWVQVEGDGQSGFSINFNITKTRPAELALLALESEAGVP 802  
 Db 787 NGLHCSLKFRPKPEINGSQFSTRVIVYQKASSTPEVMW--EVSNEPELYDATSEEREKGSV 844



```
CjSuperfamily: thrombospondin type 1 repeat homology
F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match          5.6%; Score 267.5; DB 2; Length 984;
Best Local Similarity 37.7%; Pred. No. 2.8e-11;
Matches 55; Conservative 22; Mismatches 54; Indels 15; Gaps 5;

QY      220 TCVA-----KNIVARRRSASAAVIIVVVGWGSTWTWWSVCSSCGRGWKRSRCTNPA 273
       ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      317 TCVSPYGHCSGPLRESRVCNNALCPVHGWEWSPWSLCSFTCGRGQRTTRSCT--P 374

QY      274 PLNGAFCEGQNVDHRTVSLLSYDGSWPSPKWSACGLDC---THWRSECDSPAPEN 330
       ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      375 POYGGRPCGEPETHHKPCNALCPVDGWQEWSWSQCSTCSNGTQQRSGCT--AAAH 432

QY      331 GGEECGQTDLDRNCTSDLCVHSASG 356

Db      433 GGSECRGPWAESRECYNPEC--TANG 456
```

RESULT 8  
T00028  
brain-specific angiogenesis inhibitor 3 - human  
N:Alternate names: BAI3 protein  
C:Species: Homo sapiens (man)  
C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 21-Jul-2000  
C:Accession: T00028  
R:Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.  
Cytogenet. Cell Genet. 79, 103-108, 1997  
A:Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain  
A:Reference number: Z14066; MUID:98194217; PMID:9533023  
A:Accession: T00028  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1522 <SH1>  
A:Cross-references: EMBL:AB005299; NID:g3021700; PIDN:BA25363.1; PID:g3021701  
A:Experimental source: brain

A:Gene: GUB:BAL3  
A:Cross-references: GDB:9838090; OMIM:602684  
A:Map position: 6q12-6q12  
C:Superfamily: thrombospondin type 1 repeat homology  
F:344-398/Domain: thrombospondin type 1 repeat homology <THR3>  
  
Query Match      5.6%; Score 267.5; DB 2; Length 1522;  
Best Local Similarity 37.7%; Pred. No. 5.1e-11;  
Matches 55; Conservative 22; Mismatches 54; Indels 15; Gaps 5;  
  
QY    220 TCVA-----KNIVARRRSASAAVIVVYNGWGWSTWTWSVCSSACGRGQKKSRRCTNPA 273  
     ||||| : : : : :  
Db    317 TCVSPGYTHCSGPLRESRVNCNTALCPVHGVEEWSPMSLGSFTCGRGORTTRSCT--P 374  
                : : : : :  
QY    274 PLNGGAFCBGQNVDHRTVTSSLVSVDGSWPWSKWSAAGLDLC---TWHRSRSCSDAPRN 330  
     ||||| : : : : :  
Db    375 PoyGGRPCEGPETHHKPCNTALCPVDGOWOEWSWSOCSVTCNSGTQQRSOQT--AAAH 432  
                : : : : :

Qy	331	GGECCOGTDLDTNCTSDLCVHSASG	356
		:  :  :	
		:  :  :	
Db	433	GGSECRGPWAESREYNPEC--TANG	456









QY 804 LVGPSAPK--IPFLIROKLIISLDPPCRGA--DWRTLAQKHLDSH----- 846  
 Db 829 TVDGSFVSIVPFLV-PGIRYSVEVAASTGAGEGVKSEPFQLOSHGPNVSPEDQVSLA 887  
 QY 847 --LSFFASKPSPTA-----MILNLWEARHPFNGNLSOLAASVAGLGOPDAGLFT 893  
 Db 888 QQISDVVKQPAFIAGAACWILMVFSIWLVRHKKRN--GLSSTYAGIRKVPSTFT 944

RESULT 15  
 T30805  
 dutt1 protein - mouse  
 N:Alternate names: transmembrane receptor protein Robol homolog  
 C:Species: Mus musculus (house mouse)  
 C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999  
 C:Accession: T30805  
 R:Wu, M.C.; Lowe, N.; Fordham, R.; Rabbitts, P.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: The mouse homologue of human DUTT1/H-robol gene: protein sequence and chr  
 A:Reference number: Z20879  
 A:Accession: T30805  
 A>Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1612 <WUM>  
 A:Cross-references: EMBL:Y17793; NID:el329712; PID:el329713; PIDN:CAA76850.1  
 A:Experimental source: brain  
 C:Genetics:  
 A:Gene: dutt1  
 A:Map position: 16

Query Match 4.5%; Score 217.5; DB 2; Length 1612;  
 Best Local Similarity 19.9%; Pred. No. 2.6e-07; Indels 355; Gaps 52;  
 Matches 211; Conservative 133; Mismatches 359

QY 41 DLLPHFLVPEPDVYIVKPKVLLVCKAV--PATQIEFKCNGEWVROVDHVIERS----- 92  
 Db 26 DFPFPRIVEHPSDLIVSKGEPATLCKAEGRPPTIEWYKGERV-RTDKDDPRSHMLLP 84  
 QY 93 -----TDGSSGEPTMEVRINVSROQVEKVFGLSEYWCQCVAMSSSGTTKSKQAYIR 143  
 Db 85 SGSLFFLRIVHGRKSRPDEGVYI-----CVARNYLGEAVSHNASLE 125  
 QY 144 IARLRNPFQEPILAKVSLSEOGVLPFCRPEGIPPAEVEWLRNEDLVPSLDENVYIT-R 202  
 Db 126 VAILRDDFRQNSDVNVAVGEPAVMCEQPGRHPEPTISWKDGSPLD---DKDERITIR 182  
 QY 203 EHSLVVRQARLADTANYTCVAKNIVARRSASAAVIYVNGW-----STWTMSV-- 253  
 Db 183 GCKLMITYTRKSDAGKYCVGTNMUGERESEVAELTVLERPSFVKRPSNLAVTVDDSAEF 242  
 QY 254 -CSASCGRG-----WQK-----RSRSTNPAPLNGGAPCEGQNVH 287  
 Db 243 KCEA---RGDPVTVWRKDDGELPKSRYEIRDDHTLKIRKVT--AGDMGSYTCVAENMV 297  
 QY 288 DRTVSSLVSDGS---NSPWSKWSACGLDCTHWSRECSDDAP-----RNGGEC-- 335  
 Db 298 GKAEASATITVQEPHFVVKPRDQVVALGRVTI-FQCEATGNPQPAIFWREGSQNLLPS 356  
 QY 336 ----QCTDLDRNCTSDLCVHSAGSPEDVALYVGLIAVAVCLVLLVLLVLYCRKKEGL 391  
 Db 357 YQPPQSSSRFSVSGTGLTITNVQ-RSDVGYI-----ICQTL----- 392  
 QY 392 DSDVADSSILITSGFQFVSKPSKADNPHTLTTPDLSTTTTYQGSJ---CPRQDGPSPKF 448  
 Db 393 --NVA-GSIITKAY--LEVDVIADRPFPVIRGQPVNQTVANDGTLLILSCVATGSPAPTI 447  
 QY 449 -----QLTNGHL---LSPLGG-GRHTLHSSPTSEAEFEVSRSLSTQNY 487  
 Db 448 LWRKDGVLVSTQDSRIKQLESGLQIRYAKLGDGTGRYCTASTPSGEA----- 495  
 QY 488 FRLSPRGTNMYGTNFTLGGRLMIPNTGISLLIP-----PDAIPRKIYEIVTLHKPE- 542  
 Db 496 -----TWSAY-----IEVQEFGVVQVPPRPPTDNLIPSAP-----SKPEV 530

QY 543 -DV---RPLACQOILLSPVSCGPFVLLTRPVLAMDHCGEPPSDSWSLRLKKQSCG 598  
 Db 531 TDVSKNTVTLSWQPNLNS-----GATPTSIIIEAFSHA---SGSSW-----QTAAB 573  
 QY 599 SNEQDVHLHGEAPSHLYYCOLEASACVVFTE-----Q 631  
 Db 574 NVKTETFAIKGPKNAIYFLVRAANAYGISDPQSISDPVKTDVPPTSQGVHDHKVQRE 633  
 QY 632 LGRFALVGEALSVAATAAKRLKLLFAPVACTSLEYINRVYCLHDTHTDALKEVVQLEKQLG 691  
 Db 634 LG-----NVVLHLHPTILSSSSVEV-----HWTVDQQSQYIQ-----GY 668  
 QY 692 QLIQEPRLVLFKDSYHNLRSLSDHVPSSLWKSKLL-----VSYQ--EIPFHHWNGTOR 743  
 Db 669 KILYRPS-----GASHGESEWLVEVRTPTPKNSVWIPDLRKGVNYEIKARPPFNEFOGADS 724  
 QY 744 YLHCTFTLERVSPSTSDLACKLWVQVEGDGQSFISINFN-ITKDTRFDEL-----LAL 795  
 Db 725 EIKFAKTLEE-APSAPPRS--VTVSKNDGNGTALVTWQPPEDITIONGMVQYKWKWCLGN 781  
 QY 796 ESEAGVPALVGPSAPFKIPFLIRQKIISLDDPPCR-----RGADWRTLAQKHLDS 845  
 Db 782 ETKYHINKTVDGSTFSV-----VPSLVPGIRYSVEVAASGAGPGVKSEPFQTLDS 834  
 QY 846 H-----LSFFASKPSPTA-----MILNLWEARHP--NGNLSOLA 878  
 Db 835 HGNVPSPEDQVSLAQOISDVVRQPAFIAGIAGAACWILMVFSIWLVRHKKRNLSTTYA 894  
 QY 879 A-----AVAGLGQPDAGLFTVSE 896  
 Db 895 GIRKVPSTFTTPTVYQRGGEAVSSGGRP--GULNISE 930

Search completed: October 6, 2004, 18:20:08  
 Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 15:50:42 ; Search time 19 Seconds

(without alignments)  
2463.740 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPLWPLALGIVLAWL.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280	5.8	1584	1	BAL1 HUMAN
2	273.5	5.7	1172	1	TS22 HUMAN
3	273.5	5.7	1172	1	TSP2 MOUSE
4	272.5	5.7	1077	1	SMSA MOUSE
5	270.5	5.7	1074	1	SMSA HUMAN
6	269	5.6	1170	1	TSP2 BOVIN
7	267.5	5.6	1522	1	BAL3 HUMAN
8	266.5	5.6	1173	1	TSP1 XENLA
9	258.5	5.4	1093	1	SMSB MOUSE
10	258	5.4	1572	1	BAL2 HUMAN
11	252	5.3	1093	1	SMSB HUMAN
12	248	5.2	1170	1	TSP1 MOUSE
13	246	5.1	1170	1	TSP1 HUMAN
14	243.5	5.1	1178	1	TSP2 CHICK
15	243	5.1	1170	1	TSP1 BOVIN
16	208.5	4.4	469	1	PROB HUMAN
17	205.5	4.3	470	1	PROP CAVPO
18	202.5	4.2	437	1	PROP MOUSE
19	186	3.9	1266	1	NGCA CHICK
20	181.5	3.8	867	1	SPO BOVIN
21	178.5	3.7	1223	1	AT14 HUMAN
22	175.5	3.7	1736	1	ZOL1 HUMAN
23	172	3.6	630	1	AT34 RAT
24	170.5	3.6	1745	1	ZOL1 MOUSE
25	169	3.5	934	1	CO6 HUMAN
26	167.5	3.5	905	1	AT58 MOUSE
27	166	3.5	837	1	AT94 HUMAN
28	162.5	3.4	587	1	CO8B ONCMY
29	160.5	3.4	562	1	AT15 MOUSE
30	160.5	3.4	967	1	AT51 HUMAN
31	160.5	3.4	1077	1	AT10 HUMAN
32	159.5	3.3	1224	1	AT16 HUMAN
33	156.5	3.3	860	1	AT56 HUMAN

34	156	3.3	950	1	AT15 HUMAN
35	155	3.2	997	1	AT57 HUMAN
36	154.5	3.2	1593	1	AT12 HUMAN
37	154	3.2	1213	1	AT52 MOUSE
38	153.5	3.2	1205	1	AT53 HUMAN
39	152	3.2	967	1	AT51 RAT
40	152	3.2	968	1	AT51 MOUSE
41	151.5	3.2	890	1	AT58 HUMAN
42	151.5	3.2	1095	1	AT17 HUMAN
43	151	3.2	930	1	AT55 HUMAN
44	150	3.1	1277	1	CAML FUGRU
45	149	3.1	140	1	CO8B RAT

## ALIGNMENTS

RESULT 1  
BAL1 HUMAN  
ID BAL1 HUMAN STANDARD; PRT; 1584 AA.  
AC O14514;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Brain-specific angiogenesis inhibitor 1 precursor.  
GN BAL1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=98054121; PubMed=9393972;  
RA Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,  
RA Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;  
RT "A novel brain-specific p53-target gene, BAL1, containing  
thrombospondin type 1 repeats inhibits experimental angiogenesis.";  
RL Oncogene 15:2145-2150(1997).  
RN [2]  
RP INTERACTION WITH BAP1.  
RX MEDLINE=98321173; PubMed=9647739;  
RA Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,  
Tokino T.;  
RT "Cloning and characterization of BAL-associated protein 1: a PDZ  
domain-containing protein that interacts with BAL1.";  
RL Biochem. Biophys. Res. Commun. 247:597-604(1998).  
CC -!- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN  
BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53  
SIGNAL IN SUPPRESSION OF GLOBLASTOMA. MAY FUNCTION IN CELL  
ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.  
CC -!- SUBUNIT: INTERACTS WITH BAP1.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE  
CONCENTRATED AT CELL-CELL ADHESION SITES.  
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN NO  
EXPRESSION IS OBSERVED IN SOME GLOBLASTOMA CELL LINES AND CANCER  
TISSUES.  
CC -!- INDUCTION: By p53.  
CC -!- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT  
CORNEA INDUCED BY BFGF.  
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.  
CC -!- SIMILARITY: Contains 5 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 1 GPS domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; AB005297; BAA3647.1; --











```
FT DOMAIN 784 839 TSP TYPE-1 5.
FT DOMAIN 841 896 TSP TYPE-1 6.
FT DOMAIN 897 944 TSP TYPE-1 7.
FT CARBOHYD 142 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 933 933 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 56 56 A -> V (IN REF. 2).
FT CONFLICT 149 149 A -> T (IN REF. 2).
FT CONFLICT 382 382 V -> M (IN REF. 2).
FT CONFLICT 494 494 S -> R (IN REF. 2).
SQ SEQUENCE 1074 AA; 120570 MW; E53DB763CBE29407 CRC64;

Query Match 5.7%; Score 270.5; DB 1; Length 1074;
Best Local Similarity 42.4%; Pred. No. 1.2e-12;
Matches 50; Conservative 13; Mismatches 52; Indels 3; Gaps 1;

QY 241 VNGGWTWTSVCSASCGRWKRSCTNPAPANGGAFCEGQNVHDTVSSLLYSVDG 300
DB 783 VNGASAWTSWSQCSRDCSRGIRNRKVCNNPKYGMPCLGPSLEYQFCNTLPCFVDG 842

QY 301 SWSPEKWSAGCLDGC--THWRSGCSDPAPRNGEGECQCTDLDRNCTSDLCVHSGS 355
DB 843 VWSGSPWTKSATCGGHHYMRTRSCSNPAPAGGIDICLGLHTEBALCNTQPCFWS 900

RESULT 6
TSP2_BOVIN STANDARD; PRT; 1170 AA.
AC Q95116; Q28180;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 2 precursor (Corticotropin-induced secreted protein)
DE (CISP)
GN THBS2 OR TSP2 OR TSP-2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Danik M., Chinn A., Lafeuillade M., Keramidas M., Aguesse-Germon S.,
RA Penhoat A., Chen H., Mosher D., Chambaz E.M., Feige J.J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-522 FROM N.A.
RX MEDLINE=9631130; PubMed=8698834;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RT "Opposite regulation of thrombospondin-1 and corticotropin-induced
RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic
RT hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
RN [3]
RP SEQUENCE OF 318-831 FROM N.A.
RC TISSUE=Aortic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
```

```
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFEC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X96540; CAA65385.1; -
CC EMBL; X87620; CAA60952.1; -
CC HSSP; P00740; 1EDM.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSPC.
CC InterPro; IPR003129; TSPN.
CC InterPro; IPR001007; VWF_C.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00090; tsp_1; 3.
CC Pfam; PF02412; tsp_3; 13.
CC Pfam; PF05735; TSPC; 1.
CC Pfam; PF02210; TSPN; 1.
CC Pfam; PF00093; vwc; 1.
CC PRINTS; PR01705; TSPIRREPEAT.
CC SMART; SM00181; EGF; 3.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS50026; EGF_3; 2.
CC PROSITE; PS50092; TSP1; 3.
CC PROSITE; PS01208; VWF_1; 1.
CC PROSITE; PS50184; VWF_2; 1.
CC Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
CC EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 19 215 THROMBOSPONDIN 2.
FT DOMAIN 19 232 TSP N-TERMINAL.
FT DOMAIN 318 375 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 379 429
FT DOMAIN 435 490 TSP TYPE-1 1.
FT DOMAIN 492 547 TSP TYPE-1 2.
FT DOMAIN 547 587 TSP TYPE-1 3.
FT DOMAIN 588 645 EGF-LIKE 1.
FT DOMAIN 646 690 EGF-LIKE 2.
FT DOMAIN 723 758 EGF-LIKE 3.
FT DOMAIN 759 781 TSP TYPE-3 1.
FT DOMAIN 782 817 TSP TYPE-3 2.
FT DOMAIN 818 840 TSP TYPE-3 3.
FT DOMAIN 841 878 TSP TYPE-3 4.
FT DOMAIN 879 914 TSP TYPE-3 5.
FT DOMAIN 915 950 TSP TYPE-3 6.
FT DOMAIN 951 1170 TSP TYPE-3 7.
FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 266 266 INTERCHAIN (PROBABLE).
FT DISULFID 270 270 INTERCHAIN (PROBABLE).
FT DISULFID 391 423 BY SIMILARITY.
FT DISULFID 395 428 BY SIMILARITY.
FT DISULFID 406 413 BY SIMILARITY.
FT DISULFID 447 484 BY SIMILARITY.
FT DISULFID 451 489 BY SIMILARITY.
```



```
FT DOMAIN 1045 1098 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1099 1119 6 (POTENTIAL).
FT DOMAIN 1120 1125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1126 1146 7 (POTENTIAL).
FT DOMAIN 1147 1522 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 159 CUB.
FT DOMAIN 291 343 TSP TYPE-1 1.
FT DOMAIN 345 398 TSP TYPE-1 2.
FT DOMAIN 400 453 TSP TYPE-1 3.
FT DOMAIN 455 508 TSP TYPE-1 4.
FT DOMAIN 816 868 GPS.
FT DOMAIN 942 945 POLY-THR.
FT DOMAIN 1173 1176 POLY-SER.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 779 779 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 828 828 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 937 937 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1522 AA; 171490 MW; D22DOA5D4BB62502 CRC64;

Query Match 5.6%; Score 267.5; DB 1; Length 1522;
Best Local Similarity 37.7%; Pred. No. 3.3e-12;
Matches 55; Conservative 22; Mismatches 54; Indels 15; Gaps 5;

QY 220 TCVA-----KNIVARRRSAAIVYVNGWSTWTSVCSACGRGWKRSRCTNPA 273
Db 317 TCVSPYGTCHSGPLRESRCNNATLCPVHGVMWSPWLSLCSFTGCGQRTTRSC--P 374
QY 274 PLNGAFCEGQVNDRTVSLVSDGSPSKWSACGLDC---THWRSRCSOPAPRN 330
Db 375 PQYGGPRCEGPTHHPKPCNALCPVDGQWQWSSQCSVTCNSGTFQQRSCROCT--AAAH 432
QY 331 GBECEGOTDLDTNCTSDLCVHSASG 356
Db 433 GSECEGPAWESRECVNPEC--TANG 456

RESULT 8
TSP1_XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSP1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RL Submitted (XXX-1993) to the EMBL/GenBank/DDFG databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFEC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
```

```
CC CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; L04278; -; NOT ANNOTATED_CDS.
CC DR HSSP; P00740; IEDM.
CC DR InterPro; IPR001881; EGF_Ca.
CC DR InterPro; IPR006209; EGF_like.
CC DR InterPro; IPR006210; IEGF.
CC DR InterPro; IPR000884; TSP1.
CC DR InterPro; IPR008085; TSP 1.
CC DR InterPro; IPR003367; tsp 3.
CC DR InterPro; IPR008859; TSPC.
CC DR InterPro; IPR003129; TSPN.
CC DR InterPro; IPR001007; VWF_C.
CC DR Pfam; PF00008; EGF; 2.
CC DR Pfam; PF00090; tsp 1; 3.
CC DR Pfam; PF02412; tsp 3; 13.
CC DR Pfam; PF05735; TSPC; 1.
CC DR Pfam; PF02210; TSPN; 1.
CC DR Pfam; PF00093; VWC; 1.
CC DR PRINTS; PR01705; TSP1REPEAT.
CC DR SMART; SM00181; EGF; 2.
CC DR SMART; SM00209; TSP1; 3.
CC DR SMART; SM00210; TSPN; 1.
CC DR SMART; SM00214; VWC; 1.
CC DR PROSITE; PS00022; EGF_1; FALSE_NEG.
CC DR PROSITE; PS01186; EGF_2; 1.
CC DR PROSITE; PS00026; EGF_3; 2.
CC DR PROSITE; PS00092; TSP1; 3.
CC DR PROSITE; PS01208; VWF_C; 1.
CC DR PROSITE; PS0184; VWF_C; 1.
CC KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 22
FT CHAIN 23 1173 THROMBOSPONDIN 1.
FT DOMAIN 23 235 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 23 224 TSP N-TERMINAL.
FT DOMAIN 319 376 WFEC.
FT DOMAIN 382 432 TSP TYPE-1 1.
FT DOMAIN 438 493 TSP TYPE-1 2.
FT DOMAIN 495 550 TSP TYPE-1 3.
FT DOMAIN 550 590 EGF-LIKE 1.
FT DOMAIN 591 648 EGF-LIKE 2.
FT DOMAIN 649 693 EGF-LIKE 3.
FT DOMAIN 726 761 TSP TYPE-3 1.
FT DOMAIN 762 784 TSP TYPE-3 2.
FT DOMAIN 785 820 TSP TYPE-3 3.
FT DOMAIN 821 843 TSP TYPE-3 4.
FT DOMAIN 844 881 TSP TYPE-3 5.
FT DOMAIN 882 917 TSP TYPE-3 6.
FT DOMAIN 918 953 TSP TYPE-3 7.
FT DOMAIN 954 1173 C-TERMINAL.
FT SITE 929 931 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 394 426 BY SIMILARITY.
FT DISULFID 398 431 BY SIMILARITY.
FT DISULFID 409 416 BY SIMILARITY.
FT DISULFID 450 487 BY SIMILARITY.
FT DISULFID 454 492 BY SIMILARITY.
FT DISULFID 465 477 BY SIMILARITY.
FT DISULFID 507 544 BY SIMILARITY.
FT DISULFID 511 549 BY SIMILARITY.
FT DISULFID 522 534 BY SIMILARITY.
FT DISULFID 554 565 BY SIMILARITY.
FT DISULFID 559 575 BY SIMILARITY.
FT DISULFID 578 589 BY SIMILARITY.
FT DISULFID 595 611 BY SIMILARITY.
```

FT	DISULFID	602	BY SIMILARITY.	RA	Adams R.H., Betz H., Pueschel A.W.;
FT	DISULFID	623	BY SIMILARITY.	RT	"A novel class of murine semaphorins with homology to thrombospondin
FT	DISULFID	647	BY SIMILARITY.	RT	is differentially expressed during early embryogenesis.";
FT	DISULFID	653	BY SIMILARITY.	RL	Mech. Dev. 57:33-45(1996).
FT	DISULFID	660	BY SIMILARITY.	CC	-!- FUNCTION: May act as positive axonal guidance cues.
FT	DISULFID	681	BY SIMILARITY.	CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.
FT	DISULFID	708	BY SIMILARITY.	CC	-!- TISSUE SPECIFICITY: In adult, only detected in brain.
FT	DISULFID	721	BY SIMILARITY.	CC	-!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and
FT	DISULFID	757	BY SIMILARITY.	CC	adult tissues. Its abundance decreases from E10 to birth.
FT	DISULFID	780	BY SIMILARITY.	CC	-!- SIMILARITY: Belongs to the semaphorin family.
FT	DISULFID	816	BY SIMILARITY.	CC	-!- SIMILARITY: Contains 1 Sema domain.
FT	DISULFID	839	BY SIMILARITY.	CC	-!- SIMILARITY: Contains 7 TSP type-1 domains.
FT	DISULFID	877	BY SIMILARITY.	CC	-----
FT	DISULFID	913	BY SIMILARITY.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
FT	DISULFID	949	BY SIMILARITY.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
FT	CARBOHYD	155	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	CARBOHYD	158	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	use by non-profit institutions as long as its content is in no way
FT	CARBOHYD	250	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	modified and this statement is not removed. Usage by and for commercial
FT	CARBOHYD	363	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
FT	CARBOHYD	705	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
FT	CARBOHYD	711	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-----
FT	CARBOHYD	1070	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	EMBL; X97818; CAA66398.1; .
SQ	SEQUENCE	1173	AA; 130019 MW; A9F036D516C0F24 CRC64;	DR	MGI:107555; Sema5b.
Query Match				DR	InterPro: IPR003659; Plexin-like.
Best Local Similarity 22.5%; Pred. No. 2.7e-12;				DR	InterPro: IPR002165; Plexin_repeat.
Matches 86; Conservative 54; Mismatches 147; Indels 95; Gaps 14;				DR	InterPro: IPR001627; Sema.
QY	11	LLGIVLAWLRSG-----AQQSATVANVPVGPANPDLLPHFLPEPEDIYVKNKPVLLVC- 65		DR	InterPro: IPR000884; TSP1.
Db	221	VEGTTLEAILRNKGLSMTNSVITLDNPVNGSPAIRTNYIGH-----KTKDLQAVCG 273		DR	Pfam: PF01437; PSI; 1.
QY	66	-KAVPATQIFKNGEWRVQDVHVIETSTGSSGPTMEVRINVRQVQVEKVFGLLEEYC 124		DR	Pfam: PF00090; tsp_1; 5.
Db	274	FSCDDLKLFAMKG-----LRTLVTUKDQVTKETKXNELIA 311		DR	PRINTS; PRO1705; TSP1REPEAT.
QY	125	QCVAMSSGTTKSQKAYIRIARLRNFEQ-----EPLAKEVSLQEGIVLPC 170		DR	SMART; SM00423; PSI; 1.
Db	312	QIV-----TRTPGVCLHNGVLHKNRDEWTDVSTCTECTQNSATICKVSCF---LMPC 361		DR	SMART; SM00630; Sema; 1.
QY	171	RP-----IPPAEVEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANY 219		DR	SMART; SM00209; TSP1; 4.
Db	362	TNATIPDECCPCRPWPSADDDPWSWDTPCS-----VTCGHG-IQQRGRSCDILNN 414		KW	Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
QY	220	TCVAKNIVAR-----RRSAAAVIVVNGSGMTWTEWSVCASCGRGQKRSCTN 271		KW	Developmental protein; Glycoprotein.
Db	415	PCGSSVQTRSCIQDCKRFQ-----DGNWGHSPWSSCVTCGSGQITRLCNS 467		FT	SIGNAL 1 19
QY	272	PAPLNGGAFCEQNVHRTVSSLLSVSDGWSFWKSWACGLDC---THWRSRECSDDAP 328		FT	CHAIN 20 1093
Db	468	PVPQLNGKQCEGEGRENKPCQKPCPINGQWGPWSLMDTCTVTCGGGQKREELCNNPKP 527		FT	DOMAIN 20 978
QY	329	RNGCECOGTDLDTRNCTSDLC 350		FT	DOMAIN 979 999
Db	528	QYEGKDCIGEPDTSQICNKKQDC 549		FT	DOMAIN 1000 1093
RESULT 9				FT	DOMAIN 236 518
ID	SM5B_MOUSE	STANDARD;	PRT; 1093 AA.	FT	DOMAIN 551 605
AC	Q60519;			FT	DOMAIN 606 662
DT	30-MAY-2000 (Rel. 39, Created)			FT	DOMAIN 664 713
DT	30-MAY-2000 (Rel. 39, Last sequence update)			FT	DOMAIN 721 776
DT	28-FEB-2003 (Rel. 41, Last annotation update)			FT	DOMAIN 795 850
DE	Semaphorin 5B precursor (Semaphorin G) (Sema G).			FT	DOMAIN 852 907
GN	SEMA5B OR SEMAG OR SEMG.			FT	DOMAIN 908 952
OS	Mus musculus (Mouse).			FT	DOMAIN 59 59
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			FT	CARBOHYD 95 95
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			FT	CARBOHYD 157 157
OX	NCBI_TaxID=10090;			FT	CARBOHYD 178 178
RN	[1]_TaxID=10090;			FT	CARBOHYD 287 287
RP	SEQUENCE FROM N.A.			FT	CARBOHYD 333 333
RC	STRAIN=NMRI;			FT	CARBOHYD 378 378
RC	MEDLINE=96414430; PubMed=8817451;			FT	CARBOHYD 532 532
RX				FT	CARBOHYD 539 539
				FT	CARBOHYD 547 547
				FT	CARBOHYD 602 602
				FT	CARBOHYD 728 728
				FT	CARBOHYD 944 944
				SQ	SEQUENCE 1093 AA; 120326 MW; 29E5C9B1E8108717 CRC64;
Query Match					5.4%; Score 258.5; DB 1; Length 1093;
Best Local Similarity 22.2%; Pred. No. 1e-11;					
Matches 104; Conservative 31; Mismatches 139; Indels 195; Gaps 16;					
QY	163	EQGIVLPCRPEGIPPAEVEWLRNEDLVDPDLDPNVYITREHSLVVRQARLADTANYTCV 222			
Db	737	EQFRFTCRAP-----LPDP-----HGLQFGRR---TETRTCP 767			



```
Db 442 ATCTGALDTRECSNLECPATDSKMGPMNAWSLCSKTDGTGQRRFRMCQATGTQGYPC 501
QY 337 GTDLTRCTSDLC--VHSASGPEDVAL----- 362
Db 502 GTGEEVKPCSEKRCFAPHEMCRDEYVYMLTWKAAAGEIYINCKPFPNAGSASRRCLLSA 561
QY 363 ----YVGLIAVAVCL---VLLLLVLLVYCRKKEGLDSVDADSSILTSFGQFVSIKPSKA 415
Db 562 QGVAYWGLPSPARCISHEVRYLYLSRHLAKQRMAGEGMSQVRS-LQELLARTYY 620
QY 416 DNPILLTTQPLDSTTTY-QGSLCRQDGPSPKFQLT-----NGHLLSPLOG 461
Db 621 SGDLFFSILRNVDITFRATYVPSADVDQRFQVSEFMDAENKEKWDDAQVSP--G 678
QY 462 GRHTLHSSPTSEAEFV-----SRLSTQNYFRSLPRG-----TSNMYIGTFN 504
Db 679 SVHLLR-----VSEDFHLVGDALKAFOSLSLIVTDNLVISTQRFVSAVSDITFPMRG 732
QY 505 FLG-----GRLMIPNTGISLLIP-----PDAIPRGK----- 530
Db 733 RRGKMDWRHSDELFLPKVELSLSPGPKATSGAGSPGRGRTGTPPPGCHSHQRL 792
QY 531 -----IYE-IYTLHKPEDVRLPLAGCQTLSPYVSCGPPGVLLTRPVIL 574
Db 793 PADPDESSYFVIGAVLYRTGLILPPP---RPPLAVTSRMT--VTVRPPTQPPAEPLIT 847
QY 575 A-----MDHCGPSPDSRLKKQSCGSEWQDVLHGEAPSHLYVCOLEASACYFT 629
Db 848 VELSVIINGTTDFHCASWDYS--RADASSGOWDTEHCQTLQTAHT--RCQ----- 896
QY 630 EQLGRFALVGE-----ALSVAARKLKLILFAPVACTSLEYNIRVYVCLDTHDALKEVVQ 684
Db 897 QHLSITFAVLAQPKDLTLELAGSPVPLVIGCAVSCMALLTLAIYA-----AFWRFIK 950
QY 685 LEKQGGGLIOEPRVHFKDSVHNRLSHDVPSSLKWSKLIVSQEIPFYHIMNGTORY 744
Db 951 SERSI-----ILLNFCLSI--LASNI---LILVGSRVLSKGVCWTMTAAF 990
QY 745 LHCTETLERSVSPSTDACKLWV-----LSFCVWLTEAWQSYLAVIGRMTRILVRKRFCLGWLGPALVYA 1039
Db 991 LHFFP----- 1115
QY 777 FSINFINITKTRFAELLALASEAG-VPALVGPSA-----FKIPPLIRQKI-----ISSL 824
Db 1040 VSVGTRTKGYGTSYCVLWLEGGLLYAFVGPAAVILVNMILGIIVFNKLMARDGISDK 1099
QY 825 DPPCRGAD---WRTL 837
Db 1100 SKQORAGSERCPSWASL 1115

RESULT 11
SM5B_HUMAN STANDARD; PRT; 1093 AA.
AC Q9P283;
DC TISSUE=Brain;
DT 10-OCT-2003 (Rel. 42, Created)
DI 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Semaphorin 5B precursor.
GN SEMA5B OR KIAA1445.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.;"
RL DNA Res. 7:143-150(2000).
```

CC -- FUNCTION: May act as positive axonal guidance cues (By similarity).

CC -- SUBCELLULAR LOCATION: Type I membrane protein.

CC -- SIMILARITY: Belongs to the semaphorin family.

CC -- SIMILARITY: Contains 1 Sema domain.

CC -- SIMILARITY: Contains 7 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AB040878; BAA95969.1; ALT INIT.

CC Genew; HGNC:10737; SEMA5B.

CC InterPro; IPR003659; Plexin-like.

CC InterPro; IPR002165; Plexin\_repeat.

CC InterPro; IPR001627; Sema.

CC InterPro; IPR000884; TSP1.

CC Pfam; PF01437; PSI; 1.

CC Pfam; PF01403; Sema; 1.

CC PRINTS; PR01705; TSP1REPEAT.

CC SMART; SM00423; PSI; 1.

CC SMART; SM00630; Sema; 1.

CC SMART; SM00209; TSP1; 4.

CC PROSITE; PS0092; TSP1; 5.

CC Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;

CC Developmental protein; Glycoprotein.

CC SIGNAL 1 26 POTENTIAL.

CC CHAIN 27 1093 SEMAPHORIN 5B.

CC DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 979 999 POTENTIAL.

CC DOMAIN 1000 1093 CYTOPLASMIC (POTENTIAL).

CC DOMAIN 236 518 SEMA.

CC DOMAIN 551 605 TSP TYPE-1 1.

CC DOMAIN 606 662 TSP TYPE-1 2.

CC DOMAIN 664 713 TSP TYPE-1 3.

CC DOMAIN 721 776 TSP TYPE-1 4.

CC DOMAIN 795 850 TSP TYPE-1 5.

CC DOMAIN 852 907 TSP TYPE-1 6.

CC DOMAIN 908 952 TSP TYPE-1 7.

CC CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 378 378 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 532 532 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 539 539 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 728 728 N-LINKED (GLCNAC. .) (POTENTIAL).

CC CARBOHYD 944 944 N-LINKED (GLCNAC. .) (POTENTIAL).

CC SEQUENCE 1093 AA; 119866 MW; 71DFEFB87CEAF0EF CRC64;

Query Match 5.3%; Score 252; DB 1; Length 1093;

Best Local Similarity 29.4%; Pred. No. 3.3e-11;

Matches 68; Conservative 34; Mismatches 85; Indels 44; Gaps 10;

QY 241 VNGGWTWTSVCSACGRGWKRSCTNPALINGAFCEGQNVHRTVSSLLVSDG 300

Db 851 VRGAWCWTWSVCSACGGHYORTSCTSAPSPGEDICLG--LHTEALCATQACP 908

QY 301 SWSPWKSACGLDCTHWRSECDPAPRNGEECCQGTDLDTNRCT--SDL-CVHSASGPE 358

Db 909 GWPSPSEWSKCTDDGAQSRSRHCELLP--GSSACAGNSSQRPCPYSEIPVILPASME 966

QY 359 DVALYVG-----LIATAVCLVL---LLLVLLVYCR---KKEGLSDVDADSSILTSFGQPV 408







DR PDB; ILSL; 18-DEC-02.  
 DR GlycoSuiteDB; P07996; -.  
 DR Genew; HGNC:11785; THBS1.  
 DR MIM; 180600; -.  
 DR GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.  
 DR GO; GO:0004872; F:signal transducer activity; TAS.  
 DR GO; GO:0007275; F:development; TAS.  
 DR InterPro; IPR001881; EGF Ca.  
 DR InterPro; IPR006209; EGF like.  
 DR InterPro; IPR008084; TSP1.  
 DR InterPro; IPR008085; TSP.  
 DR InterPro; IPR003367; TSP\_1.  
 DR InterPro; IPR008859; TSPC.  
 DR InterPro; IPR003129; TSPN.  
 DR InterPro; IPR001007; VWF\_C.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00090; tsp\_1; 3.  
 DR Pfam; PF02412; tse\_3; 13.  
 DR Pfam; PF05735; TSPC; 1.  
 DR Pfam; PF02210; TSPN; 1.  
 DR Pfam; PF00093; vwc; 1.  
 DR PRINTS; PR01705; TSPIREPEAT.  
 DR SMART; SM00181; EGF; 3.  
 DR SMART; SM00209; TSP1; 3.  
 DR SMART; SM00210; TSPN; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS00022; EGF 1; FALSE\_NEG.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 2.  
 DR PROSITE; PS00092; TSP1; 3.  
 DR PROSITE; PS01208; WFC\_1; 1.  
 DR PROSITE; PS0184; WFC\_2; 1.  
 DR Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;  
 KW EGF-like domain; Signal; 3D-structure.  
 FT SIGNAL 1 18  
 FT CHAIN 19 1170  
 FT DOMAIN 19 232  
 FT DOMAIN 24 221  
 FT TSP N-TERMINAL.  
 FT TSP TYPE-1 1.  
 FT TSP TYPE-1 2.  
 FT TSP TYPE-1 3.  
 FT EGF-LIKE 1.  
 FT EGF-LIKE 2.  
 FT EGF-LIKE 3.  
 FT TSP TYPE-3 1.  
 FT TSP TYPE-3 2.  
 FT TSP TYPE-3 3.  
 FT TSP TYPE-3 4.  
 FT TSP TYPE-3 5.  
 FT TSP TYPE-3 6.  
 FT TSP TYPE-3 7.  
 FT C-TERMINAL.  
 FT CELL ATTACHMENT SITE (POTENTIAL).  
 FT INTERCHAIN (PROBABLE).  
 FT DISULFID 270 270  
 FT DISULFID 274 274  
 FT DISULFID 391 423  
 FT DISULFID 395 428  
 FT DISULFID 406 413  
 FT DISULFID 447 484  
 FT DISULFID 451 489  
 FT DISULFID 462 474  
 FT DISULFID 504 541  
 FT DISULFID 508 546  
 FT DISULFID 519 531  
 FT DISULFID 551 562  
 FT DISULFID 556 572  
 FT DISULFID 575 586  
 FT DISULFID 592 608  
 FT DISULFID 599 617  
 FT DISULFID 620 644  
 FT DISULFID 650 663

BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.

FT DISULFID 657 676 BY SIMILARITY.  
 FT DISULFID 678 689 BY SIMILARITY.  
 FT DISULFID 705 713 BY SIMILARITY.  
 FT DISULFID 718 738 BY SIMILARITY.  
 FT DISULFID 754 774 BY SIMILARITY.  
 FT DISULFID 777 797 BY SIMILARITY.  
 FT DISULFID 813 833 BY SIMILARITY.  
 FT DISULFID 836 856 BY SIMILARITY.  
 FT DISULFID 874 894 BY SIMILARITY.  
 FT DISULFID 910 930 BY SIMILARITY.  
 FT DISULFID 946 1167 BY SIMILARITY.  
 FT CARBOHYD 248 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 360 C-LINKED (MAN).  
 FT CARBOHYD 385 /FTID-CAR 000205.  
 FT CARBOHYD 394 O-LINKED (FUC. . .).  
 FT CARBOHYD 438 /FTID-CAR 000206.  
 FT CARBOHYD 441 C-LINKED (MAN).  
 FT CARBOHYD 450 /FTID-CAR 000207.  
 FT CARBOHYD 498 /FTID-CAR 000208.  
 FT CARBOHYD 498 /FTID-CAR 000209.  
 FT CARBOHYD 507 /FTID-CAR 000210.  
 FT CARBOHYD 507 O-LINKED (FUC. . .).  
 FT CARBOHYD 708 /FTID-CAR 000211.  
 FT CARBOHYD 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 Query Match 5.1%; Score 246; DB 1; Length 1170;  
 Best Local Similarity 30.3%; Pred. No. 1.1e-10;  
 Matches 47; Conservative 24; Mismatches 66; Indels 18; Gaps 3;  
 QY 207 VYRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGSWGTWWSVCASAC 258  
 Db 399 IQQRGSCDNLNRCCEGSSVQTRTCHIQCCKRFXQ-----DGGWHSWSPWSSCVTC 451  
 QY 259 GRGWKRSRSCINPAPLNGCAFCEGQNVHDRTVSSLLVSDGWSWSPWSSVCAGLDC--- 315  
 Db 452 GDGVITRILCNPSFQMNKFCGEARETKACKKDACPFNGGWPWSPWDCVTCGGG 511  
 QY 316 THWRGRCSDPAPRNGGECQGTDLDTNRCTSLDC 350  
 Db 512 VOKRSRLCNPPTQFGGKDCVGDVTENQICNKQDC 546  
 RESULT 14  
 TSP2\_CHICK STANDARD; PRT; 1178 AA.  
 ID TSP2\_CHICK  
 AC P35440;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Thrombospondin 2 precursor.  
 GN THBS2 OR TSP2.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91217026; PubMed=2022631;  
 RA Lawler J., Duquette M., Ferro P.;  
 RT "Cloning and sequencing of chicken thrombospondin."  
 RL J. Biol. Chem. 266:8039-8043(1991).  
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
 cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
 laminin and type V collagen.  
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.  
 CC -!- SIMILARITY: Belongs to the thrombospondin family.  
 CC -!- SIMILARITY: Contains 1 VWF domain.



Search completed: October 6, 2004, 18:18:00  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 6, 2004, 17:50:33 ; Search time 84 Seconds

(without alignments)  
3376.796 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPGIWPALLGIVLAWL.....AVAGLQPDAGLFTVSEAE 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	4592.5	95.9	898	11	Q8K1S4	Q8K1S4 mus musculus
2	4545.5	95.0	898	11	O08721	O08721 rattus norv
3	2819.5	58.9	544	4	Q96GP4	Q96GP4 homo sapien
4	2705.5	56.5	931	11	O08747	O08747 mus musculus
5	2686	56.1	950	11	O8CD16	O8CD16 mus musculus
6	2679.5	56.0	931	13	Q7T2Z5	Q7T2Z5 gallus gall
7	2673.5	55.8	931	4	O95185	O95185 homo sapien
8	2570.5	53.7	943	13	O8UGT4	O8UGT4 xenopus lae
9	2505.5	52.3	1008	11	O8QY85	O8QY85 mus musculus
10	2499	52.2	945	11	O8K1S3	O8K1S3 mus musculus
11	2497	52.2	945	11	O8K722	O8K722 rattus norv
12	2493	52.1	945	11	Q9D398	Q9D398 mus musculus
13	2486.5	51.9	934	4	O81ZJ1	O81ZJ1 homo sapien
14	2479	51.8	945	4	O86SN3	O86SN3 mus musculus
15	2144.5	44.8	956	11	Q8K1S2	Q8K1S2 mus musculus
16	2136	44.6	948	4	Q8WVP7	Q8WVP7 homo sapien

17	1597.5	33.4	597	4	Q8IU70	Q8IU70 homo sapien
18	1442	30.1	328	11	Q80T71	Q80T71 mus musculus
19	1223.5	25.6	554	4	Q8N1Y2	Q8N1Y2 homo sapien
20	971	20.3	366	4	Q9H9F3	Q9H9F3 homo sapien
21	961.5	20.1	1072	5	Q9NBL0	Q9NBL0 drosophila
22	956.5	20.0	1072	5	Q9V7B5	Q9V7B5 drosophila
23	953.5	19.9	947	5	Q26262	Q26262 caenorhabdi
24	950.5	19.9	947	5	O44171	O44171 caenorhabdi
25	694	14.5	199	13	Q9PVD5	Q9PVD5 petromyzon
26	552.5	11.5	351	4	Q8TF26	Q8TF26 homo sapien
27	366	7.6	2673	4	Q96SC3	Q96SC3 homo sapien
28	366	7.6	5636	4	Q96RW7	Q96RW7 homo sapien
29	311.5	6.5	325	5	O81IK1	O81IK1 drosophila
30	306	6.4	518	4	O8IV45	O8IV45 homo sapien
31	273.5	5.7	1172	11	Q8CG21	Q8CG21 mus musculus
32	273.5	5.7	1172	11	Q7TMT3	Q7TMT3 mus musculus
33	271	5.7	1582	11	Q8CGM0	Q8CGM0 mus musculus
34	270.5	5.7	1081	5	Q9U631	Q9U631 mus musculus
35	269.5	5.6	1083	5	Q9VTT0	Q9VTT0 drosophila
36	269.5	5.6	1091	5	Q7YU67	Q7YU67 drosophila
37	267	5.6	1522	11	Q8OZF8	Q8OZF8 mus musculus
38	263.5	5.5	1461	5	Q8MYA8	Q8MYA8 caenorhabdi
39	258.5	5.4	1122	11	Q7TT33	Q7TT33 mus musculus
40	258	5.4	1573	4	Q8NGW8	Q8NGW8 homo sapien
41	256	5.3	478	11	Q8BVE5	Q8BVE5 mus musculus
42	254.5	5.3	1560	11	Q8CGM1	Q8CGM1 mus musculus
43	249	5.2	685	6	Q9TTS5	Q9TTS5 bos taurus
44	249	5.2	5146	6	Q8SPM4	Q8SPM4 bos taurus
45	248	5.2	1171	11	Q8CGB2	Q8CGB2 mus musculus

## ALIGNMENTS

RESULT 1

Q8K1S4 PRELIMINARY; PRT; 898 AA.

ID Q8K1S4  
AC Q8K1S4  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Netrin receptor Unc5h1.  
GN UNC5H1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Engelkamp D.;  
RT "Cloning of three mouse unc-5 genes and their expression patterns at mid-gestation.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ487852; CAD32250.1; --  
DR MGD; MGI:834682; Unc5h1.  
DR GO; GO:0004872; P:receptor activity; IEA.  
DR GO; GO:0007163; P:signal transduction; IEA.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP1.  
DR InterPro; IPR000906; ZUS.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00090; tsp\_1; 2.  
DR Pfam; PF00791; ZUS\_1.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00409; IG; 1.  
DR SMART; SM00203; TSP1; 2.  
DR SMART; SM00218; ZUS; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.



```

Db 420 LTIQPDLSLTITTTTQGSCLCPQDGPSPKQFQLSNGHLLSPLGGRHTLHSSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYIYTLH 539
Db 480 SRLSTQNYFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYIYTLH 539
QY 540 KPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLLKQSCGS 599
Db 540 KPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLLKQSCGS 599
QY 600 WEQDVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAPVA 659
Db 600 WE-DVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAPVA 658
QY 660 CTSEYINIRVYCLHDTHDALKEVVQLEKQGGQIQRVPLHFKDSYHNLRSLSHDVPS 719
Db 659 CTSEYINIRVYCLHDTHDALKEVVQLEKQGGQIQRVPLHFKDSYHNLRSLSHDVPS 718
QY 720 LWSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVQVGEQGQSF 779
Db 719 LWSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVQVGEQGQSF 778
QY 780 NFNTKDTREAEILLALESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTL 839
Db 779 NFNTKDTREAEILLALESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTL 838
QY 840 KLHLDHSLSPFASKPSPMTAMILNWEARHPFNGNLSQAAVAGLGOPDAGLFTVSEA 899
Db 839 KLHLDHSLSPFASKPSPMTAMILNWEARHPFNGNLSQAAVAGLGOPDAGLFTVSEA 898

RESULT 3
Q96GP4
ID Q96GP4 PRELIMINARY; PRT; 544 AA.
AC Q96GP4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to transmembrane receptor Unc5H1 (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009333; AAH09333.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00791; ZUS; 1.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00218; ZUS; 1.
KW Receptor; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 544 AA; 59949 MW; 350A7BA53375CCAE CRC64;

Query Match 58.9%; Score 2819; DB 4; Length 544;
Best Local Similarity 99.6%; Pred. No. 8.2e-252;
Matches 540; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 359 DVALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDVADSSILTSFGQPVSTKPSKADNP 418
Db 4 DVALYVGLIAVAVCLVLLVLLVLYVCRKKEGLSDVADSSILTSFGQPVSTKPSKADNP 63
QY 419 HLLTTPDLSL-TTTTQGSCLCPQDGPSPKQFQLTNGHLLSPLGGRHTLHSSPTSEAE 477

```

```

Db 64 HLLTTPDLSL-TTTTQGSCLCPQDGPSPKQFQLTNGHLLSPLGGRHTLHSSPTSEAE 123
QY 478 FVSRLSTQNTFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYIYTL 537
Db 124 FVSRLSTQNTFRSLPRGTSNMTYGTENFLGRLMIPNTGISLLIPDAIPRGKIYIYTL 183
QY 538 LHKPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLLKQSC 597
Db 184 LHKPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILAMDHCGEPPSPDSWSLRLLKQSC 243
QY 598 GSWEQDVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAP 657
Db 244 GSWE-DVLHGEAPSHLYYCQLEASACVYFTQGRFALVGEALSVAALKRLKLLPAP 302
QY 658 VACTSEYINIRVYCLHDTHDALKEVVQLEKQGGQIQRVPLHFKDSYHNLRSLSHDV 717
Db 303 VACTSEYINIRVYCLHDTHDALKEVVQLEKQGGQIQRVPLHFKDSYHNLRSLSHDV 362
QY 718 SSWKSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVQVGEQGQSF 777
Db 363 SSWKSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVQVGEQGQSF 422
QY 778 SINFNITKDTREAEILLALESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTL 837
Db 423 SINFNITKDTREAEILLALESEAGVPALVGPSPAFKIPFLIRQKIISLDPCCRRGADWRTL 482
QY 838 AQKLHLDHSLSPFASKPSPMTAMILNWEARHPFNGNLSQAAVAGLGOPDAGLFTVSEA 897
Db 483 AQKLHLDHSLSPFASKPSPMTAMILNWEARHPFNGNLSQAAVAGLGOPDAGLFTVSEA 542
QY 898 EC 899
Db 543 EC 544

RESULT 4
O08747
ID O08747 PRELIMINARY; PRT; 931 AA.
AC O08747;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rostral cerebellar malformation protein.
GN UNC5H3 OR RCM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57B6/SJL;
RX MEDLINE=97271898; PubMed=9126743;
RA Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B.,
RA Knowles B.B.;
RT "The mouse rostral cerebellar malformation gene encodes an UNC-5-like
RL protein."
RL Nature 386:838-842(1997).
DR EMBL; U72634; AAB54103.1; -.
DR MGD; MGI:1095412; Unc5h3.
DR GO; GO:0005886; C:plasma membrane; IC.
DR GO; GO:0005042; F:netrin receptor activity; IDA.
DR GO; GO:0005515; F:protein binding; IDA.
DR GO; GO:0007420; F:brain development; IMP.
DR GO; GO:0030334; P:regulation of cell migration; IMP.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR008844; TSP1.
DR InterPro; IPR008085; TSP1.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.

```

Db Pfam; PF00791; ZUS; 1.  
DR PRINTS; PRO1705; TSP1REPEAT.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00209; TSP1; 2.  
DR SMART; SM00218; ZUS; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00092; TSP1; 2.  
KW Immunoglobulin domain.  
SQ SEQUENCE 931 AA; 103062 MW; 8A5D951A4EECA179 CRC64;

Query Match 56.5%; Score 2705.5; DB 11; Length 931;  
Best Local Similarity 56.4%; Pred. No. 6.2e-241;  
Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;

QY 9 PALLGIVLAALRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVVIYVKNKPVLLVC 65  
Db 26 PAL--ALLSAGTGSAAQDDFFHELPEFPDPPEPLPHFLPEPEAYIVKNKPVNLYC 83  
QY 66 KAVPATQIIPKCNSEWVRQVDHVIERTSDGSSGEPTEWVRINVSROQVEKVGLEBYWCQ 125  
Db 84 KASPATQIYFKCNSEWVRQVDHVDERTSGLIVREVSIEISROQVEELFGPDYWCQ 143  
QY 126 CVAWSSGTTTQSKAYIRIARLKNKPEOPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR 185  
Db 144 CVAWSSAGTTKSKAYIRIARLKTFEQPLGKEVSLQGIIVLPCRPEGIPPAEVEWLR 203

RESULT 5  
Q8CD16 PRELIMINARY; PRT; 950 AA.  
AC Q8CD16;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE UNCS5H3.  
GN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RL Nature 420:563-573 (2002)."  
DR EMBL; AK031655; BAC27495.1; -.  
DR PIR; PT0566; PT0566.  
DR MGD; MGI:1095412; Unc5h3.  
DR GO; GO:0005886; Cytoplasmic membrane; IC.  
DR GO; GO:0005042; F-actin receptor activity; IDA.  
DR GO; GO:0005515; F-actin binding; IDA.  
DR GO; GO:0007420; P-kinase activity; IMP.  
DR GO; GO:0030334; P-regulation of cell migration; IMP.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR003599; Ig-like.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig C2.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP1.  
DR InterPro; IPR000906; ZUS.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00090; tsp\_1; 2.  
DR Pfam; PF00791; ZUS; 1.  
DR PRINTS; PRO1705; TSP1REPEAT.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00409; IG; 1.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00209; TSP1; 2.  
DR SMART; SM00218; ZUS; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00092; TSP1; 2.  
SQ SEQUENCE 950 AA; 105398 MW; 1E8FC74703351AF6 CRC64;

Query Match 56.1%; Score 2686; DB 11; Length 950;  
Best Local Similarity 55.3%; Pred. No. 4.1e-239;  
Matches 515; Conservative 153; Mismatches 214; Indels 50; Gaps 13;

QY 9 PALLGIVLAALRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVVIYVKNKPVLLVC 65  
Db 26 PAL--ALLSAGTGSAAQDDFFHELPEFPDPPEPLPHFLPEPEAYIVKNKPVNLYC 83  
QY 66 KAVPATQIIPKCNSEWVRQVDHVIERTSDGSSGEPTEWVRINVSROQVEKVGLEBYWCQ 125  
Db 84 KASPATQIYFKCNSEWVRQVDHVDERTSGLIVREVSIEISROQVEELFGPDYWCQ 143  
QY 126 CVAWSSGTTTQSKAYIRIARLKNKPEOPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR 185  
Db 144 CVAWSSAGTTKSKAYIRIARLKTFEQPLGKEVSLQGIIVLPCRPEGIPPAEVEWLR 203

DR Pfam; PF00791; ZUS; 1.  
DR PRINTS; PRO1705; TSP1REPEAT.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00209; TSP1; 2.  
DR SMART; SM00218; ZUS; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00092; TSP1; 2.  
KW Immunoglobulin domain.  
SQ SEQUENCE 931 AA; 103062 MW; 8A5D951A4EECA179 CRC64;

Query Match 56.5%; Score 2705.5; DB 11; Length 931;  
Best Local Similarity 56.4%; Pred. No. 6.2e-241;  
Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;

QY 9 PALLGIVLAALRGSGAQQA---TVANPVPGANPDLLPHFLVEPEDVVIYVKNKPVLLVC 65  
Db 26 PAL--ALLSAGTGSAAQDDFFHELPEFPDPPEPLPHFLPEPEAYIVKNKPVNLYC 83  
QY 66 KAVPATQIIPKCNSEWVRQVDHVIERTSDGSSGEPTEWVRINVSROQVEKVGLEBYWCQ 125  
Db 84 KASPATQIYFKCNSEWVRQVDHVDERTSGLIVREVSIEISROQVEELFGPDYWCQ 143  
QY 126 CVAWSSGTTTQSKAYIRIARLKNKPEOPLAKEVSLQGIIVLPCRPEGIPPAEVEWLR 185  
Db 144 CVAWSSAGTTKSKAYIRIARLKTFEQPLGKEVSLQGIIVLPCRPEGIPPAEVEWLR 203  
QY 186 NEDVPSLDPNVITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYNGW 245  
Db 204 NEDIIDPDRNYITIDNLIQKARLSTANTYTCVAKNIVARRRSASAAVIVYNGW 263  
QY 246 STWTEWSVCSACGRGWKESRCTNPAPLNGAFCEGQNVHRTVSSLLVSDGSGSPW 305  
Db 264 STWTEWSVCSNCRGQYKQTRCTNPAPLNGAFCEGQV-QKIACTTLCPPVDGRWTSW 322  
QY 306 SKWSACGLDTHWRSCSPAPRNGEECGTDLTFRCTSDLCVHSASGPDVALYVG 365  
Db 323 SKWSTCCTECHRRRECTAPAPKNGKDCDGLVLQSKNCTDGLCMAAPDSDVALYVG 382  
QY 366 L-TAVAVCLVLLLVLLVYCRKKEGLSDVADSSILTSFGQPVSIKPSKADNPHLLTIQ 424  
Db 383 IVIAVTVCLAITVVVALFVYKKNHRDFESDIIDSSALNGGFQPVNKAARQD---LLAVP 439  
QY 425 PDL-STTTTVOGSLCPQDQSPFPQTLNTHLLSPGLGRRHLHSS-----PTSEAEFFV 479  
Db 440 PDLTSAAMVGRGPVYALHD-VSDKIPMTNSPILDFLNLKIKVYNSSGAVTPQDDLAFFS 498  
QY 480 SRLS---TONVF-----RSLPRGT--SNMTYGTFFNLGGRIMTPTNGISLLIPDA 525  
Db 499 SKLSPQMTQSLLENEALNKNQSLARQTPDCTAGTFNSLGGHLLIPNSGVSLIPAGA 558  
QY 526 IPRGKIYIYTLHKPEDVRLPACQTLTSPVSCGPPGVLLTRPVILAMDHCGEPPSD 585  
Db 559 IPQGVVYMYVYVHRKENRPMEDSQTLTPVVSVCPPGALLTRPVILTLHHCADPSTE 618  
QY 586 SWSLRLKKQSCGSEWQVDHLGGEAPSHLYCYOLEASACVTFEOLGRFALVGEALSVA 645  
Db 619 DMKIQLNQAVQGWQ-EVVVVGGEENFTTPCYIQIDAEACHLTENLSTYALVQGSTTKA 677  
QY 646 AAKRLKLLFAPVACTSLEYNIRVYVCLDTHDALKEVVOLEKQGLQIQPRVLFKDS 705  
Db 678 AAKRLKLAIFGLCCSSLEYISRVCLDQDQALKEVLQLEQRMGQLLEPKALFKGS 737  
QY 706 YHNLRLSHDVPSSLWKLVSQYQIIPFYHINWCTQRYLHCTFTLERSVPSSTDACL 765  
Db 738 IHNRLSHDIAHSLWKLVSQYQIIPFYHINWCTQRYLHCTFTLERSVPSSTDACL 797  
QY 766 WYQVGEQSGSFINFTKTDRFALALALEAGVPALVGSAPKIPFLRQKLISSLD 825  
Db 798 CVRQVEGEGQIFQLNCTVSEETGIDPLDLPASTITVTGSAFSIFLPRQKLCSSLD 857  
QY 826 PCRGAQWRTLAOKLHSLSHLFFASKPSPTAMILNLWEARHFPNGMLSLAALAAVAGLG 885





```
O95185
ID O95185 PRELIMINARY; PRT; 931 AA.
AC O95185;
DT 01-WAY-1999 (TtEMBLrel. 10, Created)
DT 01-WAY-1999 (TtEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TtEMBLrel. 25, Last annotation update)
DE Transmembrane receptor UNC5C.
UNCL
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99000841; PubMed=9782087;
RA Ackerman S.L., Knowles B.B.;
RT "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.";
RL Genomics 52:205-208 (1998).
DR EMBL; AF055634; AAC67491.1; -.
DR Genew; HGNC:12569; UNC5C.
DR GO; GO:0005042; Finitrin receptor activity; TAS.
DR GO; GO:0007411; P:axon guidance; TAS.
DR GO; GO:0007420; P:brain development; TAS.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSP1; 2.
KW Immunoglobulin domain; Receptor.
SQ SEQUENCE 931 AA; 103101 MW, EFD71122C98DABB8 CRC64;

Query Match 55.8%; Score 2673.5; DB 4; Length 931;
Best Local Similarity 55.5%; Pred. No. 5.7e-238;
Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;

Qy 9 PALLGIVLAALRGSGAQQS---ATVANPVGANPDLLPHLPVPEPVYVKNKPVLLVC 65
Db 26 PAL--ALLSAGTGSAAQDDFFHELPEPFPDPEPLPHLIEPEAYIVKNKPVNLYC 83
Qy 66 KAVPATQIFPKNGEWYQVDHVIERSTGSSGEPTMEVRINVSQVQKVFGLSEYWCQ 125
Db 84 KASPATQIFKCNSEWHQKHIDVERDVSGLIVREVSIEISRQVVEELFGPEYWCQ 143
Qy 126 CVAWSSGTTKSKAYIRIARLKNFQEPLEAKVSLQGLVLCRPEPGIPPAVEWMLR 185
Db 144 CVAWSSAGTTKSKAYIRIARLKTFFQEPLEKGVSLQGLVLCRPEPGIPPAVEWMLK 203
Qy 186 NEDLADPSLDPNVYITREHSIVVRQARLADTANYTCVAKNIVARRRSASAVIVVNGW 245
Db 204 NEDIIDPVDNRNFIITDHNLIQARLSDTANTVCVAKNIVAKRSTATVIVVNGW 263
Qy 246 STWTWSEYCSACGRGQKRSCTNPAPLNGGAFCEQNVHDTVSGLLSVDPGSPW 305
Db 264 STWTWSEYCNRCGRGQKRTCTNPAPLNGGAFCEQSV-QKIACTTLCFVDGRWTPW 322
Qy 306 SKWSACGLDCTHWSRSCSDPAPNGGEECGTDLDTFNCTSDLCVHSASGPDVALYVG 365
Db 323 SKWSTCGTCTHWRRECTAPAPNGGKDCDGLVLQSKNCTDGLCMQAPDSDVALYVG 382
Qy 366 L-IAVAVCLVLLVLLVLYCRKKEGLSDVDADSSILTSFGQPVSIKPSKADNPHLLTIQ 424
```

```
Db 383 IIVAVIVCLAISSVVALFYRKNHRDFESDIIDSSALNGFGFQPVNIIKAARQD---LLAVE 439
Qy 425 PDL-STTTTYQSGSLCPROGPKFQLTNGHLLSLPLGGGRHTLHS-----SPTSEAEFV 479
Db 440 PDLTSAAMYRGVPVVALHD-VSDKIPMTNSPILDLPLNLKIKVYNTSGAVSPQDDISEFT 498
Qy 480 SRLS---TQNYF-----RSLPRGT--SNWYGTENFLGRGLMIPNTGILLIPPPDA 525
Db 499 SKLSPOMTOSLLENEALSUKNSLARQTPDCTAFGNSLGGHLIVPNSGVSLIPAGA 558
Qy 526 IPRGKIYEIYLTHKPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVILLAMDHCGEPPSPD 585
Db 559 IPQGRVYEMVTVHRKETWRPMDDSQTLITPVWSCGPPGALLTRPVLLTMHHCADENTE 618
Qy 586 SWSRLKQKSCGSEWQDVLHGEAPSHLYYCOLEASACYVTEQIGRFPALVGEALSVA 645
Db 619 DWKILLKNOAAQOWE-DVVVVGEENFTPCYIKLDAEACHILTENLSTYALVGHSTTKA 677
Qy 646 AAKRLKLLPAPVACTSLEVNIHVYCLHDTHDALKWVLEKOLGGQOLIQEPRLVHPKDS 705
Db 678 AAKRLKLLPAPVACTSLEVNIHVYCLHDTHDALKWVLEKOLGGQOLIQEPRLVHPKDS 737
Qy 706 YHNLRLSIHNDVPSLWKLVSQYQIPEFYHWNQYRVLHCTFTTLERVSPSTDLACKL 765
Db 738 THNLRLSIHNDVPSLWKLVSQYQIPEFYHWNQYRVLHCTFTTLERVSPSTDLACKL 797
Qy 766 WVVQVGGGQSPSINFTKTRFAELLALLESAGVPALVGPSPAFKIPFLIROLKLISSLD 825
Db 798 CVRQVEGEGQIFQLNCTVSEETGIDPLLDPAITTTVTGSPSAFPIPLIROLKLISSLD 857
Qy 826 PCPCRGADWETLAQKLHDSHLSFFASKPSPTAMILNWEARFPNGLSOLAAAVAGLG 885
Db 858 APQTRGHDMRLAKHLNDRILNFAFKSPTGVILDLWEAQNFPDGNLSMLAAVLEMG 917
Qy 886 QPDAGLFTVSEAE 898
Db 918 RHETVVSLLAAEGQ 930

RESULT 8
Q8JGT4 PRELIMINARY; PRT; 943 AA.
ID Q8JGT4;
AC Q8JGT4;
DT 01-OCT-2002 (TtEMBLrel. 22, Created)
DT 01-OCT-2002 (TtEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TtEMBLrel. 25, Last annotation update)
DE UNC-5 receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Anderson R.B., Holt C.E.;
RT "Expression of UNC-5 in the developing Xenopus visual system.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSP1REPEAT.
```







QY 837 LAQKLDHSLFFPASKPTAMILNWEARHPNGLNSQLAAAVAGLQPDAGLFTVSE 896  
 DB 883 LAQKLSMDRLYNFATKASPTGVLNWEARQDDGLNSLASALEEMKSEMLVAMTTD 942

QY 897 AEC 899  
 DB 943 GDC 945

RESULT 12

Q9D398 PRELIMINARY; PRT; 945 AA.  
 AC Q9D398;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE 6330415802Rik protein.  
 GN UNC5H2 OR 6330415802RIK.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK018177; BAB31108.1; -.  
 DR MGD; MGI:894703; Unc5h2.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig c2.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR InterPro; IPR000906; ZUS.  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00090; tsp.1; 2.  
 DR Pfam; PF00791; ZUS; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR SMART; SM00218; ZUS; 1.  
 DR PROSITE; PS50017; DEATH DOMAIN; 1.  
 DR PROSITE; PS50835; IG LIKE; 1.  
 DR PROSITE; PS50092; TSP1; 2.  
 KW Immunoglobulin domain  
 SQ SEQUENCE 945 AA; 103725 MW; 43D33B4524E0CBF2 CRC64;

Query Match 52.1%; Score 2493; DB 11; Length 945;  
 Best Local Similarity 52.1%; Pred. No. 3e-221;  
 Matches 497; Conservative 152; Mismatches 241; Indels 64; Gaps 18;

QY 1 MAVRCLWPAHLLGIVLAAM-----LEG--SGAQSATVANPVGPANPDLLPHFLVPEDV 53  
 DB 1 MARSQVRSALLALLLLCWDPTPTSLAGVDSAGQ---VLPDSTPSAPAEQLPFLLEPQDA 57  
 QY 54 YIVKXKPVLLVCKAVPATQIFFKNGEWVRQVDHVIERTSGSGSEPTMEVRINVSRQOV 113  
 DB 58 YIVKXKPVLLVCKAVPATQIFFKNGEWVRQVDHVIERTSGSGSEPTMEVRINVSRQOV 117  
 QY 114 EKVFGLSEYWCQCVAMSSSGTTSKQYIRIARLKNFEQELAKEVSLEQGVILCRPP 173  
 DB 118 EELFGLSEYWCQCVAMSSSGTTSKQYIRIARLKNFEQELAKEVSLEQGVILCRPP 177  
 QY 174 EGTPPAEVEMLNEDLVDPLDPNVYITREHSIVVQARLADTANTYCVAKNIVARRSA 233  
 DB 178 EGVFAEVEMLNEDVIDPAQDNFLITIDHNLIIQARLSUTANTYCVAKNIVARRST 237  
 QY 234 SAAVIVYVNGGWSMTWESVCSACSGRGWKRSCTNPAPLNGAFCEGQNVHRTVSS 293  
 DB 238 AATVIVYVNGGWSMAEWSPCNRCRGWKRTCTNPAPLNGAFCEGQ-AFQKTACT 296  
 QY 294 LLVSVDGWSPKWSACGLDCTHWSRECSDPAPRNGGECOGTDLDTNCTSDLCV-- 351  
 DB 297 TVCPVDGAWTEWSKWSACSTCAHWSRECMAPPQNGGRCGDTLLDKNCTDGLCVLN 356  
 QY 352 -----HSASGPEDVALYVGL-IYAVAVCLVLLVLLVYCRKKEGLSDVDASS-I 400  
 DB 357 QRTLNDPKHPLETSGDVALYAGLVAVFVAVVAVLMAEGLVYVRRNCRDFDTIDSSAA 416  
 QY 401 LTSGFQVSIKPSKADNPHLL--TIQDPLSTTT-TYQSGLCPRQDGPSPKQFQTNHLLS 457  
 DB 417 LTGFGFHVNFKTARPNNPOLLHPSAPDLTASAGIYRGPVVALQDS-ADKIPMTNSPLLD 475  
 QY 458 PLOGGRTLHSS-----PTSEAEFVSRLSTQNYERS----- 490  
 DB 476 PLPSLKIKVYNSSTIGSGSLADGALLGVLPCTYPGDP-SRDTHFLHRSASLSGQHL 534  
 QY 491 --LPRGTNMTYGTENFLGGRIMIPNTGISLLIPDAIPRGKIVEIYTLHKPEDVRLPL 548  
 DB 535 LGLPRDPSSVSGTFCGLGRSLPGTGVSLVPGNAIPQCKFYDLYLHINKAEST-LPL 593  
 QY 549 A-QCOTLLSPVSCGPPGVLLTTRPVILAMDHCGEPPSDSWLSRLKKQSCGSGWEQDVHL 607  
 DB 594 SEGSIQVLSVSVTCGTLLCRPVLLTVPHCAEVIAGDWIFOLKTAHQGHWE-EVVTL 652  
 QY 608 GEAPSHLYCYOLEASACVFTQGLGRFALVGSALVAAAKRLKLLLFAPVACTSLEVNI 667  
 DB 653 DEETLNTPCYQLEAKSCHILLDQGSYVFMGESYRSASVAKRLQALFALPACTSLEYS 712  
 QY 668 RYVCLDHTDALKEVQLEKQGLQIOEPRVLHFKDSYHNRLSIHDVFPSSLWKSLLV 727  
 DB 713 RYVCLDHTDALKEVQLEKQGLQIOEPRVLHFKDSYHNRLSIHDVFPSSLWKSLLV 772  
 QY 728 SYQEIPIFYHIWNGTORYLHCTFTLERVSPSTSLACKLWVQVGEQGSISFNITKDT 787  
 DB 773 KYQEIPIFYHVWNGSQRALHCTFTLERHSLASTFTCKVCVRQVGEQGIQFQLHTTLA-ET 831  
 QY 788 RFAELIALESEAG--VPALVGPSAFKIPFLIRQKIISLDPPCRGGADWTLQAKHLDS 845  
 DB 832 PAGSLDALCSAPGNATITQGYAFKIPFLIRQKICSSLDAPDSRGNDWLLQAKLSMDR 891  
 QY 846 HLSFFASKPSPTAMILNWEARHPNGLNSQLAAAVAGLQPDAGLFTVSEAEAC 899  
 DB 892 YLNVFATKASPTGVLLDWEARQDDGLNSLASALEEMKSEMLVAMATDGC 945

RESULT 13

Q8IZJ1 PRELIMINARY; PRT; 934 AA.  
 AC Q8IZJ1;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Transmembrane receptor UNC5H2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22246081; PubMed=12359238;  
 RA Komatsuzaki K., Dalvin S., Kinane T.B.;  
 RT "Modulation of G(ialpha(2)) signaling by the axonal guidance molecule  
 RT UNC5H2.";  
 RL Biochem. Biophys. Res. Commun. 297:898-905(2002).  
 DR ENBL; AY126437; AAM95701.1; -  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR00884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR InterPro; IPR00906; ZUS-  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00090; tsp\_1; 2.  
 DR PRINTS; PR00791; ZUS; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00409; IG; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR SMART; SM00218; ZUS; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS50092; TSP1; 2.  
 KW Receptor.  
 SQ SEQUENCE 934 AA; 102433 MW; 225B3F506D52B780 CRC64;  
 Query Match 51.9%; Score 2486.5; DB 4; Length 934;  
 Best Local Similarity 52.2%; Pred. No. 1.2e-220;  
 Matches 490; Conservative 149; Mismatches 255; Indels 45; Gaps 16;  
 QY 1 MAVRQLVALLGIVLAAN---LRGSGAQQ-SATVANPVGANPDLLPHLVEPEVDYIV 56  
 DB 1 MGARGARALLLALLGLLWDPLQSLQAGTDSGEVLDPSPAPAEPLPYFLQEPQDAYIV 60  
 QY 57 KKNPVLVCKVAPATQIFKCKNGEWMVROVDHVERSTDCSGEPTMEVIRNRCQVEKV 116  
 DB 61 KKNPVELCRAPATQIFKCKNGEWMVQNDHVTQEGLEATGLRVREVLQVRSRQOVEL 120  
 QY 117 FGLEEYWCOCVWSSGTTKSKAYIRIARLRNFEQELAKEVSLQGVILPCRPPEGI 176  
 DB 121 FGLEDYWCOCVWSSAGTTKSRAYVRIAYLRKNFQDEPLGKEVPLDHEVLQCRPEGV 180  
 QY 177 PPAEVEWLENEVDLPSLDPNVYITREHSLVROQLADTANYTCVAKNIVARRSASAA 236  
 DB 181 PVAEVEWLEKNEVDIPTQDTNFTLLIDHNLIIQARLSDTANYTCVAKNIVAKRSSTAT 240  
 QY 237 VIVYNGGWSWTWESVCSASGRCWKRSRCTNPAPLNGCAFCEGQVNDHRTVSSLLV 296  
 DB 241 VIVYNGGWSWAEWSPCSNRCGRWQKRTRCTNPAPLNGCAFCEGQ-AFOKTACTTIC 299  
 QY 297 SVDSGWSWMSKWSACGLDCTHWRSECDPAPRNGEECCQGTDLDTNRCTSDLVHSASG 356  
 DB 300 PVDGAWTEWSKWSACSTECARHRSRECMAPPQNGRDCSGTLLDSKNCCTDGLCQMOMLEA 359  
 QY 357 PEDVALYVGL-IYAVCLVLLVLLVILYVCRKKEGLSDVADSS-ILTSGFQFVSIKPSK 414  
 DB 360 SGDAALYAGLVVAIFVVAIVALLMAGVVVYRRNCRDFDTIDTSSAALTGGFHPVNFKTAR 419  
 QY 415 ADNPRL--TIQPDLSITT-TYQGSILCPQDGPSPKFQLTNGHLLSLPLGGRHLLHSSP 471  
 DB 420 PSNPQLLHPSVPPDUTASAGIYRGPVYALQDS-TDKIPMTNSPLLDPLPSLKVKYSSST 478

QY 472 T-----SEAEFVSRISTQNY-----FRS-----LPRCTSNMTYGTFF 503  
 DB 479 TGSFPLGADADLLGVLPPTGYPSDFARDTHFLHRSASLGSOQLLGLPRDPGSSVSGTF 538  
 QY 504 NFLGRLMIENTGISLLIPDPAIPRGKIYIYTLTHKPEDVRLPLA-GCOTLLSPIVSCG 562  
 DB 539 GCLGRLSIFGTGVSLLVPNGAIPQGFYEMYLINKAEST-LPLSGTQTVLSPSTCG 597  
 QY 563 PPGVLLTRPVILAMDHCGEPPDSWSRLRKQKCEGSEWQDVLHGBEAPSHLYYQLEA 622  
 DB 598 PTGLLLCRPVILTMPHCAEVSARDWIFQLKTAHQHWE-EVTLDEETLNTPCYQLEP 656  
 QY 623 SACVVFTEQLGREALGEALSVAARKLKLALLAPYACTSLEYNIRYVCLHDHDAKEV 682  
 DB 657 RACHILLDQGTGYFTGESYSRAVRKQLAVAPALCTSLYSRLRYCYCLEDTFVALKEV 716  
 QY 683 VOLEKOLGGOLIQEPVLFHFKDSYHNLRLSIHDPVSSLKWSKLLSVSQEIPFFYHWNQTO 742  
 DB 717 LELETLGGVYVEEPKPLMPKDSYHNLRLSLHDLPHAHWSKLLAKYQEIYFFHWSGSQ 776  
 QY 743 RYLHCTFTLERSVSPSTSDIACKLWVQVGEQSGSFNINFTKTRPAELLALASEAG-- 800  
 DB 777 KALHCTFTLERSHSLASTELTKICVRQVEGEGQIFQLHTTLA-ETPAGSLDTLCSAPGST 835  
 QY 801 VPALVGPFAKIPFLIROKIISLSDPPCRGADWRTLAOKLHLDHLSFPASKPSTAMI 860  
 DB 836 VITQLGPIYAFKIPLSIRQKICNSLDAPNSRGNDRMLAOKLSMDRYLNVFATKASPTGI 895  
 QY 861 LNLWEARHPFNGLSQLAAAVAGLQDPDAGLFTVSEAE 899  
 DB 896 LDLWEALQDDGDLNSLASALEMGKSEMLVAVATDGC 934  
 RESULT 14  
 Q86SN3 PRELIMINARY; PRT; 945 AA.  
 ID Q86SN3;  
 AC Q86SN3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE P53-regulated receptor for death and life.  
 GN P53RDL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22533857; PubMed=12598906;  
 RA Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;  
 RT "P53RDL1 regulates of p53-dependent apoptosis.";  
 RL Nat. Cell Biol. 5:216-223(2003).  
 DR EMBL; AB096256; BAC57998.1; -  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR InterPro; IPR00906; ZUS-  
 DR Pfam; PF00531; death; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00090; tsp\_1; 2.  
 DR Pfam; PF00791; ZUS; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00409; IG; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR SMART; SM00218; ZUS; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.

DR	PROSITE; PS00092; TSP1; 2.	DR	PROSITE; PS00092; TSP1; 2.
KW	Receptor.	KW	Receptor.
SQ	SEQUENCE 945 AA; 103637 MW; 56064E335F323447 CRC64;	SQ	SEQUENCE 945 AA; 103637 MW; 56064E335F323447 CRC64;
Query Match			
Best Local Similarity 51.8%; Score 2479; DB 4; Length 945;			
Matches 493; Conservative 150; Mismatches 249; Indels 60; Gaps 18;			
QY	1 MAVRPGLWALLGIVLAAN---LRGSAQO-SATVANPVPGANPDLLPHFLVEPEVDYIV 56	QY	8 WPALLGLVLAALRGS---GAQOSATVANPVPGANPDLLPHFLVEPEVDYIVKNKPVLL 63
DB	1 MGARSARGALLALLLWDPRLSQAGTDSGSEVLDPSPSAPAEPLPYFLQEPQDAYIV 60	DB	15 WLPWLGLFF--WAAGAAAARGADGSEILPDPSPA-EGTLPHFTEEPEDAYIIKSNPIAL 71
QY	57 KNPVLLVCKAVPATQIFKCKNGEWVRQVDHVIERTDGSSEPTWEVRINVSQVQVEKV 116	QY	64 VCKAVPATQIFKCKNGEWVRQVDHVIERTDGSSEPTWEVRINVSQVQVEKVGLBEYV 123
DB	61 KNPVELRCRAFPATQIFKCKNGEWVSQNDHVTQEGLEDEATGLRVREVOIEVSQVVEEL 120	DB	72 RCKARPAQIFKCKNGEWVQHNEHVSSELDSEGLKVRVFNVTQVEDFHGPEDYW 131
QY	117 FGLBEYWCOCVWSSSGTTSQKAYIRIARLRKNFEQPLAKEVSLQEGVILPCRPPEGI 176	QY	124 CQCVWSSSGTTSQKAYIRIARLRKNFEQPLAKEVSLQEGVILPCRPPEGIIPAEVW 183
DB	121 FGLBEDWCOCVWSSAGTTKSRAYIRIARLRKNFEQPLAKEVSLQEGVILPCRPPEGV 180	DB	132 CQCVWSSSGTTSQKAYIRIARLRKNFEQPLAKEVSLQEGVILPCRPPEGVPAEVEW 191
QY	177 PPAEVEWLRNEDLVDPSSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAA 236	QY	184 LRNEDLVDPSSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYNG 243
DB	181 PPAEVEWLRNEDLVDPSSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAA 240	DB	192 LKNEEIDSEODENIDTRADHNLIRQARLSDSGNYTCMAANIVAKRRSLSATVVVYNG 251
QY	237 VIYVVGWSTWTSVSCASCGRWKRSCTNPAPLNGGAPFCGQNVHDTVSSLLV 296	QY	244 GWSWTWTSVSCASCGRWKRSCTNPAPLNGGAPFCGQNVHDTVSSLLVSDGWS 303
DB	241 VIYVVGWSSWAEWSPCSNRCGRGQKTRTCTNPAPLNGGAPFCGQ-AFOKTACTTIC 299	DB	252 GWSWTWTSVSCASCGRWKRSCTNPAPLNGGAPFCGQNVHDTVSSLLVSDGWS 310
QY	297 SVNGSWPWSKASGLDCTHWRSECDPAPRNGGECGTDLDTNCTSDLCVHS--- 353	QY	304 PWSKWSACGLDCTHWRSECDPAPRNGGECGTDLDTNCTSDLCV----- 351
DB	300 PVDGAWTEWSKASGLDCTHWRSECDPAPRNGGECGTDLDTNCTSDLCVHS--- 359		
QY	354 -----ASGPEDVALVGL-IAVAVCLVLLVLLVYCRKKEGLSDVDSS-IL 401		
DB	360 LSPNSHLLERAG--DAALYAGLVVAVFVVVAILMAGVVVYRNRCDPDTDITDSSAAL 417		
QY	402 TSGFQPVSIKSKADNPHLL--TIQDPLSTTT--TYQSLCPRQDGPSPKQLNGHLLSP 458		
DB	418 TGGFHPVNFKTARPSNPQLLHPSVPPDLTASAGIYRGPVVALQDS-TDKIEMTNSPLDLP 476		
QY	459 LGGGRHLLHSSPT-----SEAEFVSRLSTQNY-----FRS----- 490		
DB	477 LPSLKVYVSSSTTSGSPGLADGADLLGVLPDGTYPSPDFARDTHFLHRSASLSGSQLLG 536		
QY	491 LPRGTSNMTYGTNFGRLMIPNTGISLIPDPAIPRGKIYETLYTLHKPEDVRLPLA- 549		
DB	537 LPRDPGSSVSGTFCGLGRLSIPGTGVSLVLPNGAIPQGFYEMYLINKAEST-LPLSE 595		
QY	550 GCQTLLSPVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRLKQSCGSEWQDVHLGE 609		
DB	596 GTQTVLSPSVTCGTGLLCPVLTTPHCAEVSARDWI FOLKTOAHQGHWE-EVVTIDE 654		
QY	610 EAPSHLYYQLEASACVYFTQGRFALVGEALSVAARLKLILLFAPVACTSLEYNIRV 669		
DB	655 ETLNTPCYQLEPRACHILLDQLTGYFTGYSRSVAKRLQAVAFALCTSLEYSLRV 714		
QY	670 YCLHDTHDALKEVQLEKQGLQIQEPRVLHFKDSYHNRLSLSHDVPSSLWKLVSIV 729		
DB	715 YCLEDTPVALKEVLELERTLGGYLVEEPKPLMFKDSYHNRLSLSHDLPFAHWRSKLAKY 774		
QY	730 QEIPFYHWTQRYLHCTFTLERSVPSSTDLACKLWVWQVGGQSPSINFNITKOTRF 789		
DB	775 QEIPFYHWTQRYLHCTFTLERSVPSSTDLACKLWVWQVGGQSPSINFNITKOTRF 833		
QY	790 AELLALAEAG--VPALVGPFAFKIPFLIRKTISSLDPPCRRGADMTFLAQKHLDSHL 847		
DB	834 GSLDTLCSAPGSTVTTQGPYAFKIPLSIRQKICNSLDAPNSRGNDRWMLAQKLSMDRYL 893		
QY	848 SFFASKPSPTAMINLWEARHPNGLSQLAAVAGLQPDAGLPTVSEAE 899		
DB	894 NYFATKASPTGVILDLWEALQDDGLNSLASALEMGKSEMLNVAATDGDG 945		



```

Db      311  VMSWSVCSPECEHLRIRECTAPPRNGKFCGLSQESENCTDGLCKLDPKPLHEIKPQ 370
QY      352  ----HSASGPEVALYVGLIYAVACVILVLLVLLVYCRKKEGLDSVADSSILTSGFQP 407
Db      371  RWSRRGIENASDIYLSGL-GAAVAVAVAVIIVGTYLRRSHSDYGVDDVIDSSALTGGFQT 429
QY      408  VSIKPSKADNPILL--TIQDLSITTYQGLCPRODGPSPKFLQNTNGHLLSPIG----- 460
Db      430  FNFKTVROGNSLLNPAWQPDILVRSYSGPIC-LQD-PLDKELMTESSLFNPLSDIKVK 487
QY      461  -----GGRH-----TLHSSPTSEAEFEVSRSLSTQNYER 489
Db      488  VQSSFMVSLGVSRAEVGHKNHSGTFPHGNNGRFGSTIHPNKT---PYIONLS----- 537
QY      490  SLPRGTSNMTYGFNFGGLMIPNTGICISLLIPDAIPRGKIYIYLTHKPEDVRLPLA 549
Db      538  SLPTRELTRTGVFGHGLGVMPNTGVSLIPHGAIPENSWEIYMSINQGP-SLQSD 596
QY      550  GCOTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDPSWSLRLKKOSCEGSWEODVLHGE 609
Db      597  GSEVLLSPEVTCGPPDMVITTPALTIHPCADVSSSEHNIHLKRTQOGKWE-EVMSVED 655
QY      610  EAPSHLYYQLEASACVYFTEQGRFALVGEALSVAAAKELKLLFAPVACTSLEYNIRV 669
Db      656  ESTS--CYCLDPFACHVLLDSFGTVALTGEPIITDCAVKQLKVAVFGCMSCNSLDYNLRV 713
QY      670  YCLHDTHDALKEVVQLEKQIQQGLIOEPVRLHFKDSYHNLRLSIHDVPSLWKSLLVSY 729
Db      714  YCVDNTPCAFQEVISDERHQQGQLLEPKLLHFKGNTFSLQSVSLDIPPLMRKPTAC 773
QY      730  QBIPEYHIWNGTORYLHCTFTLERSVSPSTDACKLWVQVEGDGQSFISINFNITKDTRF 789
Db      774  QEVPSRVWSSNRQPLHCAFSLERYTPTTQLSCKICIRQLKGHEQILQVQISILESEBE 833
QY      790  AELLALESEAGVPALVGPFAKIPFIPIROKIISSLDPPCRRGADNPTLAQKLHLDLHLSF 849
Db      834  TITFFAQEDSTFPAQTGPKAFKIPYSIRQICATFDTPNAGKDWQMLAQKNSINENLSY 893
QY      850  FASKPSPTAMILNWEARHPFNGNLSQALAAVAGLQOPDAGLFTVSEAE 898
Db      894  FATQSSPSAVILNWEARHQDQDGLDLSLACALEEIGRTHTKLSNITEPQ 942

```

Search completed: October 6, 2004, 18:19:33  
 Job time : 89 secs

mis Page blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 6, 2004, 18:20:14 ; Search time 5496 Seconds  
(without alignments)  
4884.662 Million cell updates/sec

Title: US-09-970-944-2  
Perfect score: 4787  
Sequence: 1 MAYRPLGWPALLGIVLAAML.....AVAGLGQPDAGLFTVSEAC 899

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlh  
-O=/cgn2.1/USPTO.spool/US09970944/runat\_05102004.112010.1636/app.query.fasta\_1.1095  
-DB=EST -CPMP=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-USER=US09970944 @CGN 1.1 4087 @runat\_05102004.112010.1636 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estcl:\*  
10: gb\_estc2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2686	56.1	3790	11	AK031655 Mus muscu
2	2665.5	55.7	2802	29	AY406491 Homo sapi
3	2651.5	55.4	2791	29	AY406493 Mus muscu
4	2493	52.1	3866	11	AK018177 Mus muscu
5	2269	47.4	2532	29	AY411749 Mus muscu
6	2253	47.1	2532	29	AY411747 Homo sapi
7	2169.5	45.3	2802	29	AY406492 Pan trogl
8	2055.5	42.9	2775	29	AY401471 Mus muscu
9	2050.5	42.8	2775	29	AY401469 Homo sapi
10	1890	39.5	2532	29	AY411748 Pan trogl
11	1864.5	38.9	1532	11	BC033727 Homo sapi
12	1550	32.4	2507	29	AY401470 Homo sapi
13	1373.5	28.7	1034	12	BI758231 603029876
14	1341.5	28.0	818	12	BI818609 603033362
15	1274	26.6	874	13	BQ689148 AGENCOURT
16	1268	26.5	889	13	BQ691915 AGENCOURT
17	1259.5	26.3	788	9	AI951556 wv36f04.x
18	1253	26.2	788	14	CA317532 UI-M-FW0-
19	1253	26.2	843	13	BX348193 BX348193
20	1236.5	25.8	756	13	BUE12387 UI-M-EW0-
21	1197	25.0	1175	10	BF530640 602071931
22	1192.5	24.9	751	14	CF735550 UI-M-HB0-
23	1189.5	24.8	749	14	CF735417 UI-M-HB0-
24	1186.5	24.8	723	13	BUE13458 UI-M-EW0-
25	1126.5	23.5	934	10	BF311804 UI-M-FD0-
26	1126	23.5	909	13	BX364574 BX364574
27	1122	23.4	678	14	CA749784 BX364574
28	1122.5	23.2	859	10	BF311896 UI-M-FD0-
29	1077	22.5	675	14	CA315487 UI-M-FW0-
30	1026	21.4	1040	13	BX345407 BX345407
31	1024	21.4	662	9	AL516580 AL516580
32	1022	21.3	843	12	BG913440 602811321
33	1016.5	21.2	604	10	BE314370 601147261
34	990.5	20.7	939	13	BX345406 BX345406
35	989	20.7	1201	13	BX422753 BX422753
36	934.5	19.5	720	9	AI959236 fd25g10.y
37	924	19.3	540	12	BI467799 389583.MA
38	900	18.8	824	12	BI737024 603360874
39	879	18.4	548	12	BI900830 ib81g04.Y
40	859.5	18.0	1147	13	BUE40446 AGENCOURT
41	856.5	17.9	1111	12	BG298307 602397080
42	849.5	17.7	987	13	BX452510 BX452510
43	846.5	17.7	514	13	BX522364 BX522364
44	827	17.3	815	14	CK140529 AGENCOURT
45	814.5	17.0	608	12	BM487397 pgm2n.pk0

# ALIGNMENTS

RESULT 1  
AK031655  
LOCUS  
DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length  
enriched library, clone:6030473H24 product:unc5 homolog (C.  
elegans) 3, full insert sequence.  
ACCESSION AK031655  
VERSION AK031655.1 GI:26327502  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
MEDLINE  
PUBMED 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
MEDLINE  
PUBMED 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
MEDLINE  
PUBMED 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
11076861

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3790)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. 3790  
/organism="Mus musculus"  
/mol type="mRNA"  
/strain="C57BL/6J"  
/ab\_xref="FANTOM DB-6030473H24"

FEATURES  
source

Db 607 TATCTCGGAGACATTCGAGCAGGAACCCCTTGGGAAAGGAAGTGTCTCTTGGAGCAGGAA 666  
QY 166 IleValLeuProCysArgProProGluGlyLeuProProAlaGluValGluTrpLeuArg 185  
Db 667 GTCTTACTCCAGTGTCCGCCACCTGAAGGATCCAGTGGCTGAGGTAGAAATGCTTAAG 726  
QY 186 AsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGluHisSer 205  
Db 727 AATCAAGACATAATTGATCTGCTGAAGATCGGAACCTTTATATATCTATCGATCACAA 786  
QY 206 LeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsn 225  
Db 787 CTGATCATCAAGCAAGCCCACTCTCAGATACAGCAAAATATATACCTGTGTGCAAAAT 846  
QY 226 IleValAlaArgArgSerAlaAlaValIleValTyrValAsnGlyGlyTrp 245  
Db 847 ATTGTTGCCAAGAAAGAACACCAACAGCCACTGTCTATCGTATGTTAATGTGGCTGG 906  
QY 246 SerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGlnLysArg 265  
Db 907 TCCACCTGGACAGAGTGTCTGTGTAAACAGCGCTGTGGCGAGGATATCAGAAACGC 966  
QY 266 SerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsn 285  
Db 967 ACAAGAACCTGCAACCAACCCAGCCCACTCAATGGTGGGCCCTTCTGTGAGGGGCGAGT 1026  
QY 286 ValHisAspArgThrValSerSerLeuValSerValAspGlySerTrpSerProTrp 305  
Db 1027 GTG--CAGAAATAGCATGCATCTATGTATGTCCAGTGGATGTAGTGGACTTCATGG 1083  
QY 306 SerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCysSerAsp 325  
Db 1084 AGCAAAATGGTCAACCTGTGGACTGAATGCACCCACTGGCGCAGGAGGAGTGCACGCA 1143  
QY 326 ProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCys 345  
Db 1144 CAGCCCCCAAGAACGGGGTAAGGACTGTGATGGCGCTGCTCCTCAATCAAGAACTGC 1203  
QY 346 ThrSerAspLeuCysValHis----- 352  
Db 1204 ACTGATGGGCTGTGCATGAGGAGTTCATTACCCCATTTCACTGAGCAGACACCCAG 1263  
QY 353 -----SerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeu 366  
Db 1264 AATGAATATGGATTTCTTCTGCTCCTGACTCAGATGATGTGCTCTCTACGTGGGATT 1323  
QY 367 ---IleAlaValAlaValCysLeuValLeuLeuLeuValLeuIleLeuValTyrCys 385  
Db 1324 GTGATCGCTGAACAGTCTGTCTGGCGATCACTGTTGTGGCGCTGTTGTATCGG 1383  
QY 386 ArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGlyPhe 405  
Db 1384 AAGAACCACCGTGACTTTGAGTCTGNACATCTGACTCTCTCAGCATCAATGGCGCTTT 1443  
QY 406 GlnProValSerIleLysProSerLysAlaAspAsnProHisLeuThrIleGlnPro 425  
Db 1444 CAGCCTGTGAACATCAAGGCTGCCAGCAAGAT-----CTCCTGCTGTGCCCT 1494  
QY 426 AspLeu---SerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 444  
Db 1495 GACCTCACCTCAGCTGCAGCCATGTACAGGGGACCTGTCTATGCTCTGCATCAT---GTC 1551  
QY 445 SerProLysPheGlnLeuThrAsnGlyHisLeuLeuLeuSerProLeuGlyGlyArgHis 464  
Db 1552 TCAGACAAATCCCAATGACCAACTCTCCAAATCTGGACCCCACTCCCAACTTGAAATC 1611  
QY 465 ThrLeuHisHisSerSer-----ProThrSerGluAlaGluGluValSer 480  
Db 1612 AAAGTGTAACAGCTCAGGTGCTGTCACTCCTCAGGATGACCTTGCCGAGTCTCACTCC 1671  
QY 481 ArgLeuSer-----ThrGlnAsnTyrPhe----- 488  
Db 1672 AAACGTCAACCCAGATGACCCAGTCTTGTGTAGAGAATGAGGCCCTTAACTGAGAAAC 1731

QY 489 ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPheAsnPheLeu 506  
Db 1732 CAGAGCTCCCAAGACAGACTGACCCATCTGTGCACAGCATTTGGTACTCTTCAACTCTCT 1791  
QY 507 GlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIle 526  
Db 1792 GGGGGTCACTCATCATCTTCCTAATTCAGGAGTAAAGCTTGTCTGATTCCTCGCTGGGCCAT 1851  
QY 527 ProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArgLeu 546  
Db 1852 CCTCAGGGAGAGTCTATGAATGTATGTACTGTACACAGGAAAGAAATATCAGGCC 1911  
QY 547 ProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGlyVal 566  
Db 1912 CCATGGAAAGACTCTCAGACCCCTACTTACCCCTGTGTGAGCTGTGGGCCCTCTCTGGAGCT 1971  
QY 567 LeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSer 586  
Db 1972 CTGCTGACCCGCCCTGTCTCTCATCTGTGCATCTGTGCAGACCCACACCGAGGAC 2031  
QY 587 TrpSerLeuArgLeuLysLysGlnSerCysGlySerTrpGluGlnAspValLeuHis 606  
Db 2032 TGAAGATCCAGCTCAAAACACAGGCAGTGCAGGGACAATGGGAG--GATGTTGGTG 2088  
QY 607 LeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCysTyr 626  
Db 2089 GTTGGGAGGAGAACTTCAACCCCTCTGTATCATTTCACTGATGCAGTGCAGAGGCTTCCAT 2148  
QY 627 ValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAla 646  
Db 2149 ATCCTCAGAGAACTCAGTACCTATGCCCTGTGTGGCAGTCCACCAAGCAGCT 2208  
QY 647 AlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrAsn 666  
Db 2209 GCCAAGCGTCTAAACTGGCCATCTTTGGGCCCTCTGCTGCTCTTCCCTGGAGTACAGC 2268  
QY 667 IleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGlu 686  
Db 2269 ATTAGAGTCTACTGCTGGATCACACAGGATGCCCTGAAGGAATTTTACAACTGGAG 2328  
QY 687 LysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyr 706  
Db 2329 AGGCAATGGGAGGACAGCTCTAGAAACCAAGGCTTCTCGTTTAAAGCAGCATC 2388  
QY 707 HisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSerLysLeuLeu 726  
Db 2389 CACAACCTGGCGCTGTCTATTATGACATCGCCCATTCCTCTGGAAGAGCAATGTCTG 2448  
QY 727 ValSerTyrGlnGluIleProPheTyrHisIleTyrAsnGlyThrGlnArgTyrLeuHis 746  
Db 2449 GCTAAGTATCAGGAATTCATTTTACCACATCTGGAGTGGCTCTCAAGAAACCTCCAC 2508  
QY 747 CysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrp 766  
Db 2509 TGCACCTTCACTCTGGAAGACTCAGCTTAACACAGTGGAACTGGTTTGAACCTCTGT 2568  
QY 767 ValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAsp 786  
Db 2569 GTGCGGAGGTGAAGGAGAGGCGAGATCTTCCAGCTCACTGTACTGTGTCTCAGAGNA 2628  
QY 787 ThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuValGly 806  
Db 2629 CCTACTGGCATCGATTAACCTTCTCTGTGACCTGTGTAGTACCATCACCAGTGTACCCGA 2688  
QY 807 ProSerAlaPheLysIleProPheLeuLeuArgGlnLysIleIleSerSerLysAspPro 826  
Db 2689 CCAAGTGTCTTTCAGCATTCCTCTCCCTATCCGCAAGAGCTATGCAGAGCCTGTGATGCC 2748  
QY 827 ProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHis 846  
Db 2749 CCTCAAAACAGAGGCCATGACTGGAGGATGTGGGCCCATAACTCAACTGCAGAGGTAC 2808





Best Local Similarity: 55.79%		Matches: 214	
Query Match: 55.39%		Indels: 38	
DB: 29		Gaps: 14	
US-09-970-944-2 (1-899) x AY406493 (1-2791)			
QY	9	ProAlaLeuLeuGlylleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer	28
DB	76	CGTGGGCTG-----GCCCTGCTAAGCGCGAGTGGCAGCCGGCTCCCGCTCAAGATGAT	129
QY	29	Ala-----ThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis	45
DB	130	GAATTTTTCAGAACTCCAGAAACCTTTCCATCTGACCCACTGAGCCATTTGCCACAC	189
QY	46	PheLeuValGluProGluAspValTyrIleValIleAsnLysProValLeuLeuValCys	65
DB	190	TTCTCTATTGAGCCCGAGGAAGCTTACATTGTGAAGAAACAAGCTGTGAACCTGTATTGT	249
QY	66	LysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlnTrpValArgGlnVal	85
DB	250	AAAGCCAGCCCTGCCACCCAGATCTACTTCARGTGAACAGCGAGTGGTTCATCAGAAG	309
QY	86	AspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGluValArg	105
DB	310	GACACGTAGTACGAGAGAGTAGATGAACCTCTGTCTAATTGTGAGAGAGTGAGC	369
QY	106	IleAsnValSerArgGlnGlnValGluIleValIlePheGlyLeuGluGluTyrTrpCysGln	125
DB	370	ATTGAGATTTCACGACAGAGTGGAGAACTGTTGGGCTGAAGATTACTGGTGCAG	429
QY	126	CysValAlaTrpSerSerGlyThrThyLysSerGlnLysAlaTyrIleArgIleAla	145
DB	430	TGTGTGCTGGAGCTCAGCAGGACCTACGAGAGTGGAGGAGCATACGTGCGCATGCG	489
QY	146	-----ArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu	163
DB	490	TTTCCAGATCTGGGAAGACATTCGACAGGAACCTTTGGGAAGAGTGTCTTGAG	549
QY	164	GlnGlyIleValLeuProCysArgProGluGlyIleProAlaGluValGluTrp	183
DB	550	CAGGAAGTCTTACTCCAGTGTGGCCACCTGAAGGGA-----TCCGTAGAATG	597
QY	183	pLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGln	203
DB	598	GCTAAAGAATCAACACATAATTGATCTCTGGAAGATCGGAACCTTTATATCTATCGA	657
QY	203	uHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAla	223
DB	658	TCACAACTGATCATCAGCAGCCGACTCTCAGATACAGCAATATATCTGTGTGTGC	717
QY	223	AlaAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnGln	243
DB	718	CAAAATATTGTTGCCAAGAAAAAGCACCACAGCCACTGTCATCTGTGTATGTTAATGG	777
QY	243	YGLYTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln	263
DB	778	TGGTGTGTCCACCTGGACAGAGTGTCTGTGTAAACAGCCGCTGTGGCGGAGATATCA	837
QY	263	nLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGluGln	283
DB	838	GAACGACACAGAACCTGCACCAACCCAGCCCACTCAATGGTGGGCTTCTGTGAGGG	897
QY	283	YGLAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSe	303
DB	898	GCAGAGTGTG--CAGAAAAATAGCATGCACTGATATGTCCAGTGGATGTGTAGTGGAC	954
QY	303	rProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCys	323
DB	955	TTCATGGAGCAATGTUACCTGTGGAGTGAATGCCACCACTGCGCGCAGAGGGAGTG	1014
QY	323	sSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGlyThrAspLeuAspThrAr	343
DB	1015	TACAGCACCGCCCCCAAGAACGGGGTAAGGACTGTGTATGGCTGTCTCCAATCCAA	1074

QY	343	gAsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr	363
DB	1075	GAATGCACTGATGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	1134
QY	363	rValGlyLeu-----IleAlaValAlaValCysLeuValLeuLeuLeuValLeuIle	382
DB	1135	CGTGGGATTTGATCGCTGTAAACAGTCTGTGGGATCACTGTGTGTGTGTGTGTGT	1194
QY	382	uValTyrCysArgLysGlyLeuAspSerValAlaAspSerSerIleLeuTh	402
DB	1195	TGTGTATCGGAAGAACCCCGTCACTTGTGATCTGACATCACTGATCTCTCAGCACTCA	1254
QY	402	rSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuTh	422
DB	1255	TGGCGGCTTTTACGCTGTGAACATCAAGGCTGCCAGACAGAT-----CTCCTGGC	1305
QY	422	rIleGlnProAspLeu---SerThrThrThrThrGlnGlySerLeuCysProArgGln	441
DB	1306	TGTCCTCCCTGACCTCACCTCAGCTCAGCATGTACAGGGACCTGTCTATGCTCTGCA	1365
QY	441	nAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGln	461
DB	1366	TGAT---GTCTCAGCAAAATCCCAATGACCAACTCTCCANITCTGACCCACTACCAA	1422
QY	461	YGLYArgHisThrLeuHisHisSerSer-----ProThrSerGluAlaGluGln	477
DB	1423	CTTGAATAAATAAGTGTACACAGCTCAGGTGTGTCTCTCTCAGGATGACCTTGCCGA	1482
QY	477	uPheValSerArgLeuSer-----ThrGlnAsnTyrPhe-----	488
DB	1483	GTCTCATCAAACTGTACCCAGATGACCCAGTCTCTGTCTAGAGAATGAGGCCCTTAA	1542
QY	489	-----ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPh	503
DB	1543	CCTGAAGAACCCAGAGCTCGCAAGACAGACTGACCCATCTCTGACACATTTGTACTCT	1602
QY	503	eAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProPr	523
DB	1603	CAACTCTCTTGGGGTCACTCATCTCTTCTTATCAGGAGTAAGTTGCTGATTCGCCG	1662
QY	523	oAspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAs	543
DB	1663	TGGGGCCATCTCTCAGCGGAGAGTCTATGAATGTATGTACTGTACACAGGAAGAAA	1722
QY	543	pValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyPr	563
DB	1723	TATGAGCCCTCCATGGAAGACTCTCAGACCCCTACTTACCCCTGTGTGAGCTGTGGCC	1782
QY	563	oProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSe	583
DB	1783	TCTTGGAGCTCTGTGACCCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1842
QY	583	rProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAs	603
DB	1843	CACCGAGACTTGAAGATCCAGCTCAAAACACAGGCAGTGCAGGCAATGGGAG---GA	1899
QY	603	pValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSe	623
DB	1900	TGTTGTGGTGGTGGGAGGAGAACTTCACACCCCTGTGTATTCTGAGTGGATGAGAGA	1959
QY	623	rAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSe	643
DB	1960	GGCTTGCCATCTCTCAGAGAACTTCAGTACCTATGCTCTGTGTGGCAGTCCACAC	2019
QY	643	rValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLe	663
DB	2020	CAAAAGCAGCTGCCAAGCGTCTTAAATGGCCATCTTTGGGCCCCCTCTGCTCTCTCT	2079
QY	663	uGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValVa	683
DB	2080	GGAGTACAGCAATTAGTCTACTGCTGATGACACACAGGATGCTCTGAGGAAGTTCT	2139



```

QY 693 lGlnLeuGlnLysGlnLeuGlyGlyGlnLeuLeuLeuGlnGluProArgValLeuHisPheLy 703
Db 2140 ACAACTGAGAGCAAAATGGGAGGACAGCTCCTAGAAGACCCCAAGGCTCTTCATTAA 2199
QY 703 sAspSerTyHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSe 723
Db 2200 AGGAGAGATCCAAACCTGGCCCTGTCTATTATGATCATCGCCCATTCCTCTCGAAGAG 2259
QY 723 rLysLeuLeuValSerTyGlnGlnLleProPheTyHisIleTrpAsnGlyThrGlnAr 743
Db 2260 CAATTCCTGCTGAAGTATCAGGAATTCATTTTACCATCTGGAGTGGCTCTCAAG 2319
QY 743 gTyLeuHisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCy 763
Db 2320 AAACCTCCACTGCACCTTCACTCTGGAAGACTCAGCTTAACACAGCTGGNACTGGTTG 2379
QY 763 sLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnII 783
Db 2380 CAAACTCTGTGGCGGAGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2439
QY 783 eThrLysAspThrArgPheAlaGlnLeuLeuAlaLeuGluSerGluAlaGlyValProAl 803
Db 2440 GTCAGAGAACCTTACTGGCATCGACTTACCTCTCTGAGCCCTGCTAGTACCATCACAC 2499
QY 803 aLeuValGlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSe 823
Db 2500 TGTCCGCGGACCAAGTCTTTCAGCATCTCTCCCTATCCGCGCAGAGCATGTCAGCAG 2559
QY 823 rLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLe 843
Db 2560 CTGGATGCCCTCAACAAGAGGCCATGATCGAGGATGCTGGCCCAATAAATCAACCT 2619
QY 843 uAspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLe 863
Db 2620 GGCAGGACTTCTGAATTTCTTCCACCAATCGAGCCCAATCGGCTGATCTGGATCT 2679
QY 863 uTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGl 883
Db 2680 TTGGGAAGCACAGAACTTCCAGATGGAACCTGAGCATGCTGGCAGCGCTCTCGAAGA 2739
QY 883 yLeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898
Db 2740 AATGGGAAGACATGAGACAGTGTCTTGGCAGCAGAGGACAG 2785

RESULT 4
AK018177
LOCUS
DEFINITION
Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION
AK018177.1 GI:12857775
VERSION
HTC; CAP trapper
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
99279253
MEDLINE
10349636
PUBMED
10349636
AUTHORS
Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
20499374
MEDLINE
11042159
PUBMED
11042159
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

```

```

Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3866)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 10.0 and subtraction to
Rot = 100.0. Second strand cDNA was prepared with the primer
adapter of sequence [5'
GAGAGAGAGATTCGAGTTAAATAATTAATCCCCCCCCCC 3']. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3'
end: BamHI. Host: DH10B.
Location/Qualifiers
1..3866
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:6330415E02"
/db_xref="MGI:1910526"
/db_xref="taxon:10090"
/clone="6330415E02"
/sex="male"
/tissue type="medulla oblongata"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
FEATURES
source

```

CDS

417. 3254  
/note="unnamed protein product; TRANSMEMBRANE RECEPTOR  
UNC5H2 homolog [Rattus norvegicus] (SPTT|008722, evidence:  
FASTV, 96.5%id, 100%length, match=2835)  
putative"  
/codon\_start=1  
/protein\_id="BAB31108.1"  
/db\_xref="GI:12857776"  
/translation="NRASGVRSAALLALLLLCWDPTPSLAGVDSAGQVLPDSYPSAPA  
EOLPFLPQDAIVKPKVELHCRFAFATQYFKNGEWSQNDHVTQESLDEATG  
LRVREQLVRSQVEELFGLDYMCOCVAMSSGTTKRRARIVATYLRKNFDQBP  
AKEVLDHVLVLCQPECPVAVEMLKNEVDIDPAQDTNELLTDHNLIRQARLS  
DYATYCAKNIIVAKRSTAAIVIVYNGWSMAWSPSCNRCGRGQKRETCNP  
APLNGAFCEGAQFQACTTTCVVDGAWTEWSKMSACSTECAHWSRRCMAFPQNG  
GRDCSTLLDNCSTDTGLVNRQTLNDPKSHPLETSGDVVALYGLVAVFVVAVIM  
ARGVTVYRNCDFDFTDIDSSAALTGGFHPNFKTARENFLQLLHPSAPPDLTASAG  
IYRGVVALQDSADKIPMTNSPLLDPLPSIKI KVNYSSTIGSGSLDAGDLGLVLP  
GTYGDFSDTHPLHRSASLQHLGLPRDSSVSTGFCGLGRLSLPGTGVSL  
VNGAIPQKFDVLYLHINKASTLPLSGSQTVLSPVTCGPTGLLCPVVLTVPH  
CAEVIAGDWIFOLKTAHQHMEVVTLDDETLNFCYQLKAKSCHILLDQLGSYVF  
MGEYSRVAKRQLQAFAPALCTSEYSLRVCTEDYVALKEVLELERTGLGGLVE  
EPKLLFKDSYKHLRLSLHDIPAHWRSLKAKYQIPFYHWNQSQRALHCTFTLER  
HSLASTFECKVCVROVEGQIFOLHHTLAETPAGSLDALCSAPGNATITOLGPYAF  
KPLSRQKICSLDAPDSRGNDRLLAOKLSMDRYLNYPATKASPTGVILDWEARQ  
QDDGDLNLSALAEEMKSEMLVAMATGDC"

ORIGIN

Alignment Scores:  
Pred. No.: 9, 95e-211 Length: 3866  
Score: 2493.00 Matches: 497  
Percent Similarity: 68.03% Conservative: 152  
Best Local Similarity: 52.10% Mismatches: 241  
Query Match: 52.08% Indels: 64  
DB: 11 Gaps: 18

US-09-970-944-2 (1-899) x AK018177 (1-3866)

Qy 1 MetAlaValArgProGlyLeuTrpProAlaLeuLeuGlyLeValLeuAlaIaTrp--- 19  
Db 417 ATGAGGCGCGAGCGGGTGGCGGAGCGCGTGTCTGCGCGTCTTGTGGGAT 476  
Qy 20 -----LeuArgGly-----SerGlyAlaGlnGlnSerAlaThrValAlaAsn 33  
Db 477 CGACACCGAGCGCTAGCGCGGTGACTCTGCTGCGCAG-----GTGCTCCAGAC 527  
Qy 34 ProValProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGluAspVal 53  
Db 528 TCCTACCATCAGCCCTCGCGAGCAGTGCCTGACTTCTTATTGGAGCCACAGAGCGCC 587  
Qy 54 TyrIleValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThrGlnIle 73  
Db 588 TACATCGTAAGAACCAAGCCAGTGGAACTGCACCTGCAGAGCTTCCCTGCCACAGATC 647  
Qy 74 PhePheLysCysAsnGlyGluTrpValArgGlnValAspHisValIleGluArgSerThr 93  
Db 648 TACTTCAAGTGTATGGCGAGTGGTTCGCCAGATGACCCAGTGCACAGAGAGCGCTG 707  
Qy 94 AspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArgGlnGlnVal 113  
Db 708 GATGAGGCCACAGCGCTTGGCGGTGCGAGAGTGCAGAGTGCAGAGTGTCTACCGCGCAGCAAGTG 767  
Qy 114 GluLysValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTrpSerSerSerGly 133  
Db 768 GAGGAACCTCTTGGGCTCGAGAGTACTGGTGCCAGTGGCGTGGCGTGGAGCTCTTCGGGA 827  
Qy 134 ThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGln 153  
Db 828 ACTACCAAGAGTGGCGAGCGCTTACATCCGATGCTGCTACTTGGCGAAGAACTTTGACCAG 887  
Qy 154 GluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArgProPro 173  
Db 888 GAGCCTCTGGCCAGAGAGGTACCTTGGATCATGAGTCTCTTCTGCTGCGCGCCACCG 947

Qy 174 GluGlyIleProProAlaGluValGluTrpLeuArgGlnAsnLeuValAspProSer 193  
Db 948 GAGGAGAGTGCCTGTGGCTGAGGTGGAATGCTCAAGATGAAGATGTCTATTACCCCGCT 1007  
Qy 194 LeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeu 213  
Db 1008 CAGGACACTAATCTCTGCTCACCATTGACCAACCTCATCTCCGCGAGCGCGCTC 1067  
Qy 214 AlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAla 233  
Db 1068 TCAGACACCGCCAACTACACCTGTGTGGCAGAATATCTGTGGCAAGCGCCGAGACACC 1127  
Qy 234 SerAlaAlaValIleValTyrValAsnGlyGlyTrpSerThrTrpThrGluTrpSerVal 253  
Db 1128 GGGGCCACAGTCATCGTCTATGTGAATGAGGCTGGTCCAGCTGGCAGAGTGGTCAACC 1187  
Qy 254 CysSerAlaSerCysGlyArgGlyTrpGlnLysArgSerArgSerCysThrAsnProAla 273  
Db 1188 TGTTCCAATCGCTGTGGCGAGCTGGCAGACGCTACTCGGACCTGCACCAATCCAGCC 1247  
Qy 274 ProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSer 293  
Db 1248 CCACCTCAATGGAGCGCCCTTCTGTGAGGAGCAG---GCCTTCCAGAGACAGCTTGCACC 1304  
Qy 294 LeuLeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCysGlyLeu 313  
Db 1305 ACCGTGTCGCCAGTGGATGGAGCGTGGACGAGTGGAGCAAGTGTCTGCTCGACGACA 1364  
Qy 314 AspCysThrHisTrpArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGlu 333  
Db 1365 GAGTGTGGCAGCTGGCGGAGCGCGGAGTGCATGGCACCGCCACCCAGAGAGGAGCGCT 1424  
Qy 334 GluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal----- 351  
Db 1425 GACTGAGCGGAGCAGCTACTTGTACTCAAGAACTGCACTGATGGGTGTGGTGTGAAT 1484  
Qy 352 -----HisSerAlaSerGlyProGluAspValAlaLeu 362  
Db 1485 CAGAGAACTCTAAACGACCCCTAAAGCCACCCCTGGAGACATCGGAGAGATGTGCACATG 1544  
Qy 363 TyrValGlyLeu-----IleAlaValAlaValCysLeuValLeuLeuValLeuIle 381  
Db 1545 TAGCGAGCGCTTGTGGTGGCGCTCTTGTGTGGTGTAGCGGTCTCTCATGGCGGAGGAGTG 1604  
Qy 382 LeuValTyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSer---Ile 400  
Db 1605 ATCGTATACCGAGAAACTCGCGGAGTTCGACACGACATCACCGACTCTCTCTCGCGCC 1664  
Qy 401 LeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeu 420  
Db 1665 CTCACCTGTGGCTTCCACCCCTGTCAACTTCAAGACTGCAAGCGCCCAACACCCGAGCTC 1724  
Qy 421 Leu-----ThrIleGlnProAspLeuSerThrThrThr---ThrTyrGlnGlySerLeu 437  
Db 1725 CTGACCCGCTCCCGCCCTCCAGACCTAACCGCCAGTGTGCTGCATCTACCGCGGCGCTGTG 1784  
Qy 438 CysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSer 457  
Db 1785 TATGCCCTGCAGAGTCC---GCCGACAGATCCCATGACTAATTCGCCCTCTGTGAT 1841  
Qy 458 ProLeuGlyGlyArgHisThrLeuHisHisSerSer-----ProThrSerGluAlaGluPhe 470  
Db 1842 CCCCTGCCAGCCCTCAAGATCAAGTCTATAACTCCAGACCATCGGTCTTGGGTCTTGGC 1901  
Qy 471 -----ProThrSerGluAlaGluPhe 478  
Db 1902 CTGGCTGATGAGCGGACCTCTCTGGGTGTCTCCCGCGGCGCAGCTACCCAGCGGATTTTC 1961  
Qy 479 ValSerArgLeuSerThrGlnAsnTyrPheArgSer----- 490  
Db 1962 ---TCCCGGAGACCCATTTCTGACCTGCGCAGTGCAGCCTTGGTTGCCAGCACCTC 2018  
Qy 491 -----LeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPheLeuGlyGly 508



Db	193	gatcaggctccttcgacgtgcccacccagcagggagctgctgctgagcggaa	252	Db	1264	ctggcaggtccagcagcttgggttccagcagcagcctcctgggctacctggcagcccccagcagc	1323
Qy	183	trpleuargasnclupleuValaspProSerLeuaspProasnValtyrIlethrArg	202	Qy	498	metthrTyrglyThrPheasnPhelLeuGlyArgLeuMetIleProasnThrGlyIle	517
Db	253	tgctcaagaatgaagatgtcatgtgaccccgctcagacacactnaacttctgtctacacatt	312	Db	1324	agtgtcagcggcagccttgggttgcctgggaggaagcctgagcctcccccgcacaggggtc	1383
Qy	203	gluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaasnTyrrThrCysVal	222	Qy	518	serLeuLeuIleleuProaspAlaIleProArgGlyLysIleTyrgluIleTyrrLeuThr	537
Db	313	gaccacaacctcatcctgcgcagcagcgccctctcagacacagcgcccaactacacctgtgtg	372	Db	1384	agccttcttggtaccaaatagagccattcccccagggcgaagtcttatgacctgtatctacat	1443
Qy	223	AlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrrValasn	242	Qy	538	LeuHisLysProgluAspValArgLeuProLeuAla---GlyCysGlnThrLeuLeuSer	556
Db	373	gccagaatatctgtgcccagcagcgccagcaccagcgccacagctcatcgctatgtgaat	432	Db	1444	atcaacaagcgccgaagacc---ctccacattcagaagggttcccagacagcagatttgagc	1500
Qy	243	glyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTip	262	Qy	557	ProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMet	576
Db	433	ggaggctgggtccagctggcgagagtgctcaccctgttccaatcgctgtgcccagagctgg	492	Db	1501	ccctcggtgacctgtggcgccacagcgcctactcctgtgcccgcctgtgctgctcaccgtg	1560
Qy	263	glnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyAlaPheCysGlu	282	Qy	577	AspHisCysGlyGluProSerProaspSerTipSerLeuArgLeuLysLysGlnSerCys	596
Db	493	cagaagcgtaactcggaacctcaccacactccagccccactcaatgagggcgcttctgtgag	552	Db	1561	ccccactgtgctgaagctcatcgctggagactggatctttcagctcaagaccagggcccat	1620
Qy	283	GlyGlnAsnValHisaspArgThrValSerSerLeuLeuValSerValaspGlySerTip	302	Qy	597	gluGlySerTrpGluGlnaspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrr	616
Db	553	ggacag---gcccttcagaaagacagcttgacaccgctgcccaggtggagcggtgg	609	Db	1621	cagggccactggcag---gaggtggtgaccccttgcatgagcagacccctcaacacacccctgc	1677
Qy	303	SerProTrpSerLysTrpSerAlaCysGlyLeuaspCysThrHisTrpArgSerArgGlu	322	Qy	617	TyrCysGlnLeuGluAlaSerAlaCysTyrrValPheThrGluGlnLeuGlyArgPheAla	636
Db	610	accaggtggagcagagtgcttgccttgacgacagagtgctggcactggcgacgcccggag	669	Db	1678	tactgacagctgagggctaaagctcctccacatcctgctggaccagctgggttacctagcta	1737
Qy	323	CysSeraspProAlaProArgasnGlyGlyGluGluCysGlnGlyThrAspLeuAspThr	342	Qy	637	LeuValGlyGluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuLeuPheAla	656
Db	670	TGCATGCGACCGCCACCCAGACGGAGCGCGTCACTGCAGCGGACGCTACTTGACTCC	729	Db	1738	ttcatggcgagcttctactctctgctgcatgacggcgtccagctggccatcttgcgc	1797
Qy	343	ArgasnCysThrSeraspLeuCysVal-----	351	Qy	657	ProValAlaCysThrSerLeuGluTyrrAsnIleArgValTyrrCysLeuHisAspThrHis	676
Db	730	AAGAACTGCACATGATGGGCTGTGGCTGTGAATCAGAGAACTTAAACGACCCCTAAAGC	789	Db	1798	ccagccctctgcacctcctcgagatagcctcagggctctactctctggaggacacacct	1857
Qy	352	HisSerAlaSerGlyProgluAspValAlaLeuTyrrValGlyLeu---IleAlaValAla	370	Qy	677	AspAlaLeuLysGluValValGlnLeuLysGlnLeuGlyGlyGlnLeuIleGlnGlu	696
Db	790	cacccctggagacatcggagatgtggcactgtacgacgcttgggtggcgctctt	849	Db	1858	gtgacctgaaggaggtcctggagctggagaggaactctgggtggcttacttgggtggagag	1917
Qy	371	ValCysLeuValLeuLeuLeuValLeuLeuLeuValTyrrCysArgLysLysGluGly	390	Qy	697	ProArgValLeuHisPheLysAspSerTyrrHisAsnLeuArgLeuSerIleHisAspVal	716
Db	850	GTGGTGGTAGGGTCTCATGCGCGGTGGAGTGTATACCGGAGAACTGCGCGGAC	909	Db	1918	cccgaagccttgccttttaggcagattaccacacactacgacctctccctccatcacatc	1977
Qy	391	LeuaspSeraspValAlaAspSerSer---IleLeuThrSerGlyPheGlnProValSer	409	Qy	717	ProSerSerLeuTrpLysSerLysLeuLeuValSerTyrrGlnGluIleProPheTyrrHis	736
Db	910	ttcacacagggacatcacccgactcctctgcggccctctactgtgtggtctccacccctgtcaac	969	Db	1978	ccccatgccccactggagagcnaactacttgcccaagtagccaggagattcccttctaccac	2037
Qy	410	IleLysProSerLysAlaAspAsnProHisLeuLeu---ThrIleGlnProaspLeu	427	Qy	737	IleTrpAsnGlyThrGlnArgTyrrLeuHisCysThrPheThrLeuGluArgValSerPro	756
Db	970	ttcagactgaagcccaacacccggagctctctgcacccgctgcgccccctccagaccta	1029	Db	2038	gtctggaaatggaccagagagccctgcactgcacttttccacctggagagcagatagcctg	2097
Qy	428	SerThrThrThr---ThrTyrrGlnGlySerLeuCysProArgGlnAspGlyProSerPro	446	Qy	757	SerThrSeraspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer	776
Db	1030	ACGGCCAGTGTGCGCATCTACCGCGGCGCTGTGTATGCCCTGCGAGACTCC---GCCGAC	1086	Db	2098	gcctccagcgagttcacctgaagctctgctggcgaggtcgaaggaggagggcagatt	2157
Qy	447	LysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGlyGlyArgHisThrLeu	466	Qy	777	PheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGlu	796
Db	1087	AAGATCCCATGACTAAITCGGCCCTCTGCTGGATCCCTGCCCGACGCTCAAGATCAAGTGC	1146	Db	2158	ttcagctgcacacaaactgtggcc---gagacgcttgcctgctgctcctgctgctctctgc	2214
Qy	467	HisHisSerSer-----	470	Qy	797	SerGluAlaGly-----ValProAlaLeuValGlyProSerAlaPheLysIleProPhe	814
Db	1147	TATACTCCAGCACCACCTCGGTTCTGGGTTCTGGCTGGCTGTGATGGAGCCGACCTGGGT	1206	Db	2215	tctgccccgggaatggcctaccacccagctggagacctatgcttccagataccctctg	2274
Qy	471	-----ProThrSerGluAlaGluPheValSerArgLeuSerThrGlnAsnTyrr	487	Qy	815	LeuIleArgGlnLysIleIleSerSerLeuaspProProCysArgArgGlyAlaaspTrp	834
Db	1207	gtcctccccggcgacgtaccagcgatttc---tccccgggacacccatwtctctgtgcac	1263	Db	2275	tccatcccccaaaagatctgcagcagcctggacgcccccaactcccgggcggaacgactgg	2334
Qy	488	PheArgSer-----LeuProArgGlyThrSerAsn	497	Qy	835	ArgThrLeuAlaGlnLysLeuHisLeuaspSerHisLeuSerPhePheAlaSerLysPro	854
				Db	2335	agcctgttggcgacagaagctgtccatgagcggtagcttacttgcacacaaagct	2394



```

488 ----PheArgSer-----LeuProArgGlyThrSer 496
1261 CACCTGGCAGCCAGCCTCGGTCCAGACGCTCTTGGGGCTTGGCCCGAGACCCAGGG 1320
497 AsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGly 516
1321 AGCAGCGTCAGCGGCACCTTGTGCTGGTGGGAGGCTCAGCATCCCGCCGACAGGG 1380
517 IleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeu 536
1381 GTACGTTGTGTGTCCTATCCAGGCAATCCCGAGGCAAGTCTACGAGATGTATCTA 1440
537 ThrLeuHisLysProGluAspValArgLeuProLeuAla---GlyCysGlnThrLeuLeu 555
1441 CTCATCAACAGCAGCAAGATACC---CTCCCGCTTTCAGAAAGGACCCACAGCATATG 1497
556 SerProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAla 575
1498 AGCCCTCGGTGACCTGTGACCCACAGGCTCTCTGTGTGGCCCGCTCATCTTCACC 1557
576 MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSer 595
1558 ATGCCCACTGTGCGAAGTCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1617
596 CysGluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu 615
1618 CACCAGGCGCACTGGGAG---GAGTGTGTGATCCCTGGATGAGGAGACCTGACACACCC 1674
616 TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPhe 635
1675 TGCTACTGCCAGCTGGAGCCAGGCGCTGTACATCTCTGTGGACACAGCTGGGACCTAC 1734
636 AlaLeuValGlyGluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPhe 655
1735 GTGTTACGGCGAGTCTATTCCTCGCTCAGCAGTCAAGCGCTCCAGCTGCGCGTCTTC 1794
656 AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr 675
1795 GCCCGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1854
676 HisAspAlaLeuLysGluValGlnLeuGlyGlnLeuGlyGlyGlnLeuIleGln 695
1855 CCTGTAGCAGTGAAGAGTCTGAGCTGGAGCGGACTCTGGCGGAGTACTTGGTGAG 1914
696 GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp 715
1915 GAGCGGAACCGCTAATGTTCAAGGACAGATTACACACACCTGCGCTCTGCTGCTGCTG 1974
716 ValProSerSerLeuTrpLysSerLysLeuValSerTyrGlnGluIleProPheTyr 735
1975 CTCGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2034
736 HisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer 755
2035 CACATTTGGAGTGGCAGCAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2094
756 ProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGln 775
2095 TTGGCTCCACAGAGCTCACCTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2154
776 SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeu 795
2155 ATATTCAGCTGCATACCACTCTGGCA---GAGACACCTGCTGGCTGCTGCTGCTGCTG 2211
796 GluSerGluAlaGly-----ValProAlaLeuValGlyProSerAlaPheLysIlePro 813
2212 TGCTTGCCCTGGGAGCACTGTACACCCAGCTGGGACCTTATGCTTCAAGATCCCA 2271
814 PheLeuIleArgGlnLysIleLysSerSerLeuAspProProCysArgArgGlyAlaAsp 833
2272 CTGTCCATCGCCAGAGATATGCAACAGCTAGATGCCCCCACTGCTGCTGCTGCTGCTG 2331

```

834 TrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys 853  
2332 TGGCGATGTAGCAGAGCTCTATGAGCCGCTACCTGAATTACTTTCACCAAAA 2391  
854 ProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn 873  
2392 GCGAGCCCAACGGTGTGATCTCTGGAGCTCTGCGAGCAGTCTGCGAGCAGCAGTGGGAC 2451  
874 LeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr 893  
2452 CTCACAGCTGCGAGTGCCTTGGAGGAGATGGCAAGAGTGCATGCTGGTGGCTGTG 2511  
894 ValSerGluAlaGluCys 899  
2512 GCCACCGCAGGACTGC 2529

RESULT 7  
AY406492 2802 bp DNA linear GSS 15-DEC-2003  
LOCUS Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AY406492.1 GI:39762466  
VERSION AY406492.1  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM  
REFERENCE 1 (bases 1 to 2802)  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
JOURNAL Gene trios  
PUBMED Science 302 (5652), 1960-1963 (2003)  
REFERENCE 2 (bases 1 to 2802)  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..2802  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>2802  
/gene="UNC5C"  
/locus\_tag="HCM2575"

Alignment Scores:  
Pred. No.: 4,656-182 Length: 2802  
Score: 2169.50 Matches: 438  
Percent Similarity: 62.95% Conservative: 138  
Best Local Similarity: 47.87% Mismatches: 306  
Query Match: 45.32% Indels: 33  
DB: 29 Gaps: 13

US-09-970-944-2 (1-899) x AY406492 (1-2802)

9 ProAlaLeuLeuGlyIleValLeuAlaAlaTrpLeuArgGlySerGlyAlaGlnGlnSer 28  
76 CTGTCCTG-----GCCCTGCTCAGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 129  
29 -----AlaThrValAlaAsnProValProGlyAlaAsnProAspLeuLeuProHis 45

Db 130 GACITTTTTCATGAACCTCCAGAAACTTTTCTTCGGATCCACCTGACGCTCTGCCACAT 189  
Qy 46 PheLeuValGluProGluAspValTyrIleValLysAsnLysProValLeuLeuValCys 65  
Db 190 TTTCTATTGAACTGAGGAGCTTATATTGTGAGATAGACCTGTGAACTGTACTGT 249  
Qy 66 LysAlaValPro-----AlaThrGlnIlePhePheLysCysAsnGlyGluTrpValArg 83  
Db 250 AAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 309  
Qy 84 GlnValAspHisValIleGluArgSerThrAspGlySerSerGlyGluProThrMetGlu 103  
Db 310 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 369  
Qy 104 ValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGluGluTyrTrp 123  
Db 370 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 429  
Qy 124 CysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIleArg 143  
Db 430 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 489  
Qy 144 IleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeuGlu 163  
Db 490 NNNNNNNNNNNCGAAGACATTTGACGAGAACCCCTAGGAAAGGAAGTCTCTTTGGAA 549  
Qy 164 GlnGlyIleValLeuProCysArgProGluGlyIleProAlaGluValGluTrp 183  
Db 550 CAGGAAGTCTTACTCCAGTCTCGGCCACTGAGGGATCCAGTGGCTGAGNNNNNNNN 609  
Qy 184 LeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArgGlu 203  
Db 610 NNGAAAAATGAAGACATAATGTGATCCGTTGGAAGATCGGAATTTTATATATTATGAT 669  
Qy 204 HisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysValAla 223  
Db 670 CACAACCTCATATAAGCAGGCGCCGACTCTGTACTGCAATATATACCTGTGTGCC 729  
Qy 224 LysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsnGly 243  
Db 730 AAAAAATTTGTCGAAGAGGAAAAAGTACAACTGCCACTGTCTATGTCTATGTCAACGGT 789  
Qy 244 GlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrpGln 263  
Db 790 GGTGCTGCTCACCCTGGACGGAGTGTCTGTGTGTAACAGCGCTGTGGACGAGGTATCAG 849  
Qy 264 LysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyClyAlaPheCysGluGly 283  
Db 850 AAACGTACAGGACTTGTACCAACCCGGCACCACCTCAATGGGGGTGCCCTTCTGTGAAGG 909  
Qy 284 GlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrpSer 303  
Db 910 CAGAGTGTG---CAGAAATAGCTGTACTACGTTATGCCAGTGGATGGCAGGTGAGC 966  
Qy 304 ProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGluCys 323  
Db 967 CCATGGACAAAGTGTACTTGTGGAACTGAGTGACCCACTGGCGCAGGAGGAGTGC 1026  
Qy 324 SerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeuAspThrArg 343  
Db 1027 ACGGCGCCAGCCCCCAAGAAATGAGCAAGGACTCGATGGCTCGTCTGTGAATCCCAAG 1086  
Qy 344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363  
Db 1087 AACTGCACGTATGGGCTTGTGATGACAGCTGCTCTGATTGATGATGATGCTCTCTAT 1146  
Qy 364 ValGlyLeu---IleAlaValAlaValCysLeuValLeuLeuValLeuLeu 382  
Db 1147 GTTGGGATTTGATAGCAGTGTGTTTGGCTGGCGATCTCTGTAGTTGGCCCTGTTT 1206  
Qy 383 ValTyrCysArgLysGluClyLeuAspSerAspValAlaAspSerSerIleLeuThr 402  
Db 1207 GTGTATCGGAAGAAATCATGCTGACTTTTGAGTCAGATATATTGACTCTTTCGGCACTCAAT 1266

Qy 403 SerGlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeuThr 422  
Db 1267 GGGGGCTTTTACGCTGTGAACATCAAGCAGCAAGCAAGAT-----CTGCTGGCT 1317  
Qy 423 IleGlnProAspLeu---SerThrThrThrThrTyrGlnGlySerLeuCysProArgGln 441  
Db 1318 GTACCCCGACACCTCAGTCTAGCTGAGCCGATGACAGAGGAGCCTGTCTATGCCCTGCAT 1377  
Qy 442 AspGlyProSerProLysPheGlnLeuLeuThrAsnGlyHisLeuLeuSerProLeuGlyGly 461  
Db 1378 GAC---GTCTCAGACAAAAATCCCAATGACCACTCTCCAATTTCTGGATCCACTGCCCAAC 1434  
Qy 462 GlyArgHisThrLeuHisHisSerSer-----ProThrSerGluAlaGluGlu 477  
Db 1435 CTGAAATCAAGTGTACACACCTCAGTCTGCTCACCCCCCAAGATGACCTCTCTGAG 1494  
Qy 478 PheValSerArgLeuSer-----ThrGlnAsnTyrPhe----- 488  
Db 1495 TTTACGTCGAAGCTGTCTCCCTCAGATGCCAGTCATTTGTTGGAGAAATGAAGCCCTCAGC 1554  
Qy 489 -----ArgSerLeuProArgGlyThr-----SerAsnMetThrTyrGlyThrPhe 503  
Db 1555 CTGAAGAACCAAGAGTCTAGCAAGCAGACTGATCCATCTCTACCGCATTTGGCAGGTTT 1614  
Qy 504 AsnPheLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSerLeuLeuIleProPro 523  
Db 1615 NNNNGCTGGNNNNNCACTTATTTGTTCCCAATTCAGGAGTCACTGTCTGATTCCTCGCT 1674  
Qy 524 AspAlaIleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAsp 543  
Db 1675 GGGGCCATTTCCCAAGGAGAGTCTAGAAATGTATGTACTGTACACAGGAAGAAACT 1734  
Qy 544 ValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyPro 563  
Db 1735 ATGAGGCCACCATGATGACTCTCAGACACTTTTTCACCCCTCTGTGAGCTGTGGGCC 1794  
Qy 564 ProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSer 583  
Db 1795 CAGGAGCTCTGCTCACCGCCCGCTGCTCTTACTATGATCACTCGCGCAGACCCCAAT 1854  
Qy 584 ProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGlySerTrpGluGlnAsp 603  
Db 1855 ACCGAGCATGGAAATATCTGCTCAGAACACGAGGACAGGAGCAGTGGGAG---GAT 1911  
Qy 604 ValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSer 623  
Db 1912 GTGGTGTGCTCGGGAGGAAAACTTCACCCCCCTGTCTACATTCAGCTGGATGCAGAG 1971  
Qy 624 AlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSer 643  
Db 1972 GCTGCCACATCTCTCACAGAGAACCTCAGCCCTCGTAGGACATTTCCACCACC 2031  
Qy 644 ValAlaAlaLysArgLeuLysLeuLeuLeuPheAlaProValAlaCysThrSerLeu 663  
Db 2032 AAGCGGCTGCGAAGCGCTCAAGTGGCCATCTTTGGGCCCTGTGTCTGCTCTCGCTG 2091  
Qy 664 GluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValVal 683  
Db 2092 GAGTACAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGNAGGAAATTTTA 2151  
Qy 684 GlnLeuGluLysGlnLeuGlyGlnLeuIleGlnGluProArgValLeuHisPheLys 703  
Db 2152 CATCTTGAGACAGATGGAGGACAGCTCTCTAGAAACCTTAAGGCTCTTCATTTTAA 2211  
Qy 704 AspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTrpLysSer 723  
Db 2212 GGCAGCCCAACCTCGGCTGTCAATTCAGATATCGCCCATTCCTCTCTGGAAGAGC 2271  
Qy 724 LysLeuLeuValSerTrpGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArg 743  
Db 2272 AAATTGCTGGCTAAATATACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2331



QY	744	TyrLeuHisCysThrPheThrIleuGluArgValSerProSerThrSerAspLeuAlaCys	763		
Db	2332	AACTGCATGCACCTTCACTCTGGAAGATTAGCCTGAACACATGGAGCTGGTTG	2391		
QY	764	LysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIle	783		
Db	2392	AAACTCTGTGTGGCGGAGGTGGGAAGAGAGAGGCGAGATCTTCAGCTCACTGACCGGTG	2451		
QY	784	ThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAla	803		
Db	2452	TCAGAGGAACCTACTTGGCATCGATTGCCGCTGTGGATCTCTGCAACCATCACCACG	2511		
QY	804	LeuValGlyProSerAlaPheLysIleProPheIleuIleArgGlnLysIleIleSerSer	823		
Db	2512	GTACAGGGGGCCAGTGTCTTACAGCATCCCTCTCCCTATCCGCGAGAAGCTCTGTAGCAGC	2571		
QY	824	LeuAspProProCysArgArgGlyValaAspTrpArgThrLeuAlaGlnLysLeuHisLeu	843		
Db	2572	CTGGATGCCCCCCAGACAGAGGGCCATGACTGGAGGATGCTGGGCCATAAGCNNNNNNN	2631		
QY	844	AspSerHisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuAsnLeu	863		
Db	2632	NNNNNGTACTTGAATTAATCTTGGCCACCAATCCAGCCCAACTGCGCTTGGATCTT	2691		
QY	864	TrpGluAlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGly	883		
Db	2692	TGGGAAGCAGCAAGACTTCCAGATGGAAACCTTGAGCATGCTGGCAGCTGCTTGAAGAA	2751		
QY	884	LeuGlyGlnProAspAlaGlyLeuPheThrValSerGluAlaGlu	898		
Db	2752	ATGGGAAGACATGAACCGTGTCTTACGACGAGAGGGCAG	2796		
RESULT 8	AY401471	2775 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY401471	Mus musculus HCM0901 gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY401471	genomic survey sequence.			
ACCESSION	AY401471.1	GI:39757460			
VERSION	GSS.				
KEYWORDS					
SOURCE	Mus musculus (house mouse)				
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES					
source	1. .2775				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10090"				
gene	<1_->2775				
ORIGIN	/locus_tag="HCM0901"				
Alignment Scores:					



QY 344 AsnCysThrSerAspLeuCysValHisSerAlaSerGlyProGluAspValAlaLeuTyr 363  
 DB 1003 AACTGCACAGATGCTCTCGCATCTAGGCAATTCAGAGTGCAGGCAATGCTTTGTAC 1062  
 QY 364 ValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLeuLeuLeuLeuVal 383  
 DB 1063 TCCGGCTTG---GGTGACGCGGTGGTAGCTGTTGGGTCTGGTCAATGGTGTCAACCTC 1119  
 QY 384 TyrCysArgLysLysGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSer 403  
 DB 1120 TACAGCGGACCCACAGTACGAGTACGGAGTGTGATGCTCTTCGCAATTCACAGGT 1179  
 QY 404 GlyPheGlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu---- 421  
 DB 1180 GGCTTCCAGACCTTCAACTTCAAAACAGTTCGTCAAGGAAATTCCTACTACTGAATCCT 1239  
 QY 422 ThrIleGlnProAspLeuSerThrThrThrThrThrThrThrThrThrThrThrThr 441  
 DB 1240 GCCATGCGACGAGTCTGACTGTGAGCGGACATACAGCGGCGCCATCTGT---CTGCAG 1296  
 QY 442 AspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly--- 460  
 DB 1297 GAT---CGCTCGACAGAGGCTCATACAGAGTCTCACTCTTCAATCTTTATCTGAC 1353  
 QY 460 ----- 460  
 DB 1354 ATCAAGTCAAGTCCAGAGTCAATTTATGTTTCCCTGGGAGTGTCTGAGAGCGGAG 1413  
 QY 461 -----GlyGlyArgHis----- 465  
 DB 1414 TACATGGCAAGATCACTCGGGACTTCCCCCATGGAAACAAACCGCGGATTCAGTACA 1473  
 QY 466 LeuHisHisSerSerProThrSerGluAlaGluPheValSerArgLeuSerThrGln 485  
 DB 1474 ATACATCCCAACAAACG-----CCGTACATCAAAATCTGTCA----- 1515  
 QY 486 AsnTyrPheArgSerLeuProArgGlyThrSerAsnMetThrTyrGlyThrPheAsnPhe 505  
 DB 1516 -----TCACTGCCGCAAGGACAGAGCTGAGGACAACCTGGTGTCTTTGGCCAT 1563  
 QY 506 LeuGlyGlyArgMetIleProAsnThrGlyIleSerLeuLeuIleProProAspAla 525  
 DB 1564 TTAGGGGAGCGCTAGTATGTCACAAATACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1620  
 QY 526 IleProArgGlyLysIleTyrGluIleTyrLeuThrLeuHisLysProGluAspValArg 545  
 DB 1621 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1680  
 QY 546 LeuProLeuAlaGlyCysGlnThrLeuLeuSerProIleValSerCysGlyProProGly 565  
 DB 1681 CTGCAGTCAGATGGATCTGAGGTCTCTCTGAGTCTGAAGTCACTGTGGGCTCCAGAT 1740  
 QY 566 ValLeuLeuThrArgProValIleLeuAlaMetAspHisCysGlyGluProSerProAsp 585  
 DB 1741 ATGCTGTGCACAACTCCCTTTCGCTGACCATCCCTCACTGTGCAGACGTCAGTTCAGAG 1800  
 QY 586 SerTyrPheLeuArgLeuLysGlnSerCysGlySerThrGluGlnAspValLeu 605  
 DB 1801 CACTGGAACATTCACCTGAAGAGAGGACCCAGAGGGAAGTGGAG---GAAGGTATG 1857  
 QY 606 HisLeuGlyGluGluAlaProSerHisLeuTyrTyrCysGlnLeuGluAlaSerAlaCys 625  
 DB 1858 TCAGTGGAGGATGAGTCCACATCC-----TGTTACTGCTCTTGATCCCTTTCCTGT 1911  
 QY 626 TyrValPheThrGluGlnLeuGlyArgPheAlaLeuValGluAlaLeuSerValAla 645  
 DB 1912 CACGTGCTCTGTATGCTTTGGAACGATGTCCTCACAGGAGGCAATCACAGACTGT 1971  
 QY 646 AlaAlaLysArgLeuLysLeuLeuPheAlaProValAlaCysThrSerLeuGluTyr 665  
 DB 1972 GCCGTAAAGCAGCTCAAGTGGCTGTGTGGTGTGCTGCTGCTCACTTAACTCTTGGATTAC 2031  
 QY 666 AsnIleArgValTyrCysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeu 685

DB 2032 AACCTGAGAGTTTACTGTGTGCAAACTACTCCTTGTGCAATTCAGGAAGTATTTCTGAC 2091  
 QY 686 GluLysGlnLeuGlyGlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSer 705  
 DB 2092 GAACGGCACCAAGCGCGCCAGCTCTCTGGAAGAACCAAGCTGCTGCATTTCAAGGGGAAT 2151  
 QY 706 TyrHisAsnLeuArgLeuSerIleHisAspValProSerSerLeuTyrLysSerLysLeu 725  
 DB 2152 ACATTCACTCTCCAGGTCTCTGTCTTGACATCCCTCCTCTCTGAGGATCAAAACG 2211  
 QY 726 LeuValSerTyrGlnGluIleProPheTyrHisIleTrpAsnGlyThrGlnArgTyrLeu 745  
 DB 2212 TTCACTGCATGCCAGGAAGTCCCTTCTCCGAGTGTGGACAGCAACCCGACCCCTGT 2271  
 QY 746 HisCysThrPheThrLeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeu 765  
 DB 2272 CACTGTGCTTCTCCCTGGAGCGCTACACGCCACACACCCAGCTGTCTGCAAAATC 2331  
 QY 766 TrpValTrpGlnValGluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLys 785  
 DB 2332 TGCATTCCGGAGCTCAAGGCCATGAACAGATCTCCCAAGTCGACAGATCCATCCTAGAG 2391  
 QY 786 AspThrArgPheAlaGluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuVal 805  
 DB 2392 AGTGAAGCAGAAACCATCACTTCTTCGCAAGAGGACAGTACCTTCTCTGCACAGACT 2451  
 QY 806 GlyProSerAlaPheLysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAsp 825  
 DB 2452 GGCCCAAAAGCCTTCAAAATTCCTATTCATCAGACAGAGGATCTCGCCACGTTTGAT 2511  
 QY 826 ProProCysArgGlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSer 845  
 DB 2512 ACCCAAAATGCCAAGGCAAGGACTGGCAGATCTTAGCAGAGAAGAACAGCATCAATAGG 2571  
 QY 846 HisLeuSerPhePheAlaSerLysProSerProThrAlaMetIleLeuLeuLeuTrpGlu 865  
 DB 2572 AATTATCATATTCCTACCCAAAGCAGCCCTTCGCTGTCTATTGAACTCTGGGAA 2631  
 QY 866 AlaArgHisPheProAsnGlyAsnLeuSerGlnLeuAlaAlaAlaValAlaGlyLeuGly 885  
 DB 2632 GCTCGTCATCAGCAGGATGGGACCTCGACTCGCTGCTGCTGCTCTGAGGAGATGGG 2691  
 QY 886 GlnProAspAlaGlyLeuPheThrValSerGluAlaGlu 898  
 DB 2692 AGGACACACAGAAACTATCCAACTATCAGACAGCCACAG 2730

RESULT 9  
 AY401469  
 LOCUS  
 DEFINITION  
 Homo sapiens HCM0901 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 ACCESSION  
 AY401469  
 VERSION  
 AY401469.1 GI:39757458  
 KEYWORDS  
 GSS.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 2775)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 JOURNAL  
 PUBMED  
 14671302  
 REFERENCE  
 2 (bases 1 to 2775)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.

Direct Submission		Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
COMMENT		This sequence was made by sequencing genomic exons and ordering them based on alignment.	
FEATURES		Location/Qualifiers	
source		1..2775	
gene		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		<1..>2775	
		/locus_tag="HCM0901"	
ORIGIN			
Alignment Scores:			
aligned. No.:	1.99e-171	Length:	2775
Score:	2050.50	Matches:	407
Percent Similarity:	61.28%	Conservative:	158
Best Local Similarity:	44.14%	Mismatches:	306
Query Match:	42.83%	Indels:	52
DB:	29	Gaps:	12
US-09-970-944-2 (1-899) x AY401469 (1-2775)			
QY	12	LeuGlyIleValLeuAlaAlaThrLeuArgGlySerGlyAlaGlnGlnSerAlaThrVal	31
Db	14	GTAGGCTGTGTTCTGATGCTCTTTTTCAT-TTCACAGGAACCTGCAATGCGAAGCCCTT	72
QY	32	AlaAsnProValProGlyAlaAsnProAspLeuLeuProHisPheLeuValGluProGlu	51
Db	73	CCCGAATCCATCCCATCAGCT---CCTGGACACTGCTCTATTTCATAGAGGCCAGAT	129
QY	52	AspValTyrIleValLysAsnLysProValLeuLeuValCysLysAlaValProAlaThr	71
Db	130	GATGCTTATATTACAGAGCAACCTATTGCACTCAGGTGCAAGCGAGGCCAGCCATG	189
QY	72	GlnIlePhePheLysCysAsnGlyGluTyrValArgGlnValAspHisValIleGluArg	91
Db	190	CAGATATTCTTCAATGCAACGGCGAGTGGGTCCATCAGAACGACGACGCTCTCTGAAGAG	249
QY	92	SerThrAspGlySerSerGlyGluProThrMetGluValArgIleAsnValSerArgGln	111
Db	250	ACTCTGGACAGAGCTCAGGTTTGAAGTCCCGAAGTGTTCATCAATGTTACTAGGCAA	309
QY	112	GlnValGluLysValPheGlyLeuGluGluTyrTyrCysGlnCysValAlaTyrSerSer	131
Db	310	CAGGTGGAGGACTTCCATGGCCCGAGGACTATTGTCAGTGTGGCGTGGAGCCAC	369
QY	132	SerGlyThrThrLysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPhe	151
Db	370	CTGGGTACTCTCAAGACAGGAGGCGCTCTGTGGCATAGCCTATTATTACGGAAAACTTT	429
QY	152	GluGlnGluProLeuAlaLysGluValSerLeuGluGlnGlyIleValLeuProCysArg	171
Db	430	GAACAAAGACCCACAGAGGAGGAGTCCCATTTGAGGCATGATTGACTGCATCCGC	489
QY	172	ProProGluGlyIleProProAlaGluValGluTyrLeuArgAsnGluAspLeuValAsp	191
Db	490	CCACCAGAGGAGTCCCTGCTGCCGAGGTGGAAATGGCTGAAAAATGAAGAGCCCATTC	549
QY	192	ProSerLeuAspProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAla	211
Db	550	TCTGAACAGACGAGAACATTGACACCGAGGCTGACCAATACCTGATCATCAGGACGCA	609
QY	212	ArgLeuAlaAspThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArg	231
Db	610	CGGCTCTCGGACTCAGAAATTACCTTCATGCGCAGCCACACATCGTGGCTAAGAGGAGA	669
QY	232	SerAlaSerAlaAlaValIleValTyrValAsnGlyTyrSerThrTyrThrGluTyr	251
Db	670	AGCCTGTCGGCCACTGTTTGGTCTACGTGAATGGAGGCTGTCTCTCTGGACAGAGTGG	729
QY	252	SerValCysSerAlaSerCysGlyArgGlyTyrGlnLysArgSerArgSerCysThrAsn	271
Db	730	TCAGCTGCAATGTTTCGCTGTGTAGAGGATGGCAGAAACGTTTCCGGACCTGCACCAAC	789
QY	272	ProAlaProLeuAsnGlyGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrVal	291
Db	790	CCAGCTCTCTCAATGTTGGGGCCCTTTTGTGAGGGAATGTCACTGCAGAAATAACCTGC	849
QY	292	SerSerLeuLeuValSerValAspGlySerTyrPsrProTyrPsrSerLysTyrSerAlaCys	311
Db	850	ACTTCT---CTTTGCTCTGTGATGGAGCTGGGAAGTGTGGAGCGAATGGTCCGCTTGC	906
QY	312	GlyLeuAspCysThrHisTyrArgSerArgGluCysSerAspProAlaProArgAsnGly	331
Db	907	AGTCAGAGAGTGTGAACATTTGCGGATCGCGGAGTGCACAGCACACCCCGAGAAATGGG	966
QY	332	GlyGluGluCysGlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysVal	351
Db	967	GGCAAAATCTGTGAAGTCTTAAGCCAGGAATCTGAAACTGCACAGATGGTCTTTGCATC	1026
QY	352	HisSerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeuIleAlaValAlaVal	371
Db	1027	CTAGGCAATTGAGAAATGCCAGCGACATTTCTTTGACTCGGGCTTG---GGTGTGCTC	1083
QY	372	CysLeuValLeuLeuLeuValLeuIleLeuValTyrCysArgLysLysGluGlyLeu	391
Db	1084	GTGCGCTTGCAGCTCCTGCTCATTTGTTGTCACCTTTACAGCGAGGACGAGTGACTAT	1143
QY	392	AspSerAspValAlaAspSerSerIleLeuThrSerGlyPheGlnProValSerIleLys	411
Db	1144	GGCGTGACGCTCATGACTCTTCTGCAATTCACAGGTGGCTTCCAGACCTTCACTTCAA	1203
QY	412	ProSerLysAlaAspAsnProHisLeuLeu-----ThrIleGlnProAspLeuSerThr	429
Db	1204	ACAGTCCGTCAAGTAACTCCCTGCTCTGAAATCTGCCATCGACGACAGATCTCACAGTG	1263
QY	430	ThrThrThrTyrGlnGlySerLeuCysProArgGlnAspGlyProSerProLysPheGln	449
Db	1264	AGCGGACATACAGCGACCATCTGT---CTGCAGGAC---CCTCTGGACAAGGAGCTC	1317
QY	450	LeuThrAsnGlyHisLeuLeuSerProLeuGly-----	460
Db	1318	ATGACAGAGCTCCTCACTCTTTAACCTTTTTCGGACATCAAGTGAAGTCCAGAGCTCG	1377
QY	460	-----	460
Db	1378	TTCAATGTTTCCCTGGGAGTGTCTGAGAGCTGAGTACCACGCGCAAGAAATCATTCACAGG	1437
QY	461	-----GlyGlyArgHisThrLeuHisSerSerProThrSerGluAlaGlu	476
Db	1438	ACTTTTCCCATGTGAAACACACAGCTTTAGTACAATGTCATCCCGAGAAATAAATG---	1494
QY	477	GluPheValSerArgLeuSerThrGlnAsnTyrPheArgSerLeuProArgGlyThrSer	496
Db	1495	CCCTACATCCAAAATCTGTCA-----TCACTCCCAAGGACAGAA	1536
QY	497	AsnMetThrTyrGlyThrPheAsnPheLeuGlyGlyArgLeuMetIleProAsnThrGly	516
Db	1537	CTGAGGACAACTGTGTCTTTGGCCATTAGGGGGCGCTTAGTAATGCCAAATACAGNN	1596
QY	517	IleSerLeuLeuIleProProAspAlaIleProArgGlyLysIleTyrGluIleTyrLeu	536
Db	1597	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1653
QY	537	ThrLeuHisLysProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSer	556
Db	1654	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1713
QY	557	ProIleValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMet	576
Db	1714	CCTGAAGTCCACTGTGTCTCTCCAGACATGATGCTCAACACTCCCTTTGCAATGACCATC	1773
QY	577	AspHisCysGlyGluProSerProAspSerThrSerLeuArgLeuLysLysGlnSerCys	596

```
Db 1774 CGCAGTGTGCAGATGTGAGTTCAGTTCAGCAATGGAATATCCATTTAAAGAGAGACACAG 1833
Qy 597 GluGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyr 616
Db 1834 CAGGCAAAATGGAG---GAAGTGTATGTCAGTGGAGATGATCATCATCC-----TGT 1884
Qy 617 TyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAla 636
Db 1885 TACTGCTTTTGGACCCCTTGGCGTGTATGTCCTGGACAGCTTTGGGACCTATGCG 1944
Qy 637 LeuValGlyGluAlaLeuSerValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 656
Db 1945 CTCAGTGAGAGCCAAATCACAGATGTGCCGTGAAGCAACTGAAGGTGGCGTTTGGC 2004
Qy 657 ProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHis 676
Db 2005 TGCATGCTCTGTAACCTCCCTGGATTACACTTGAGAGTTTACTGTGTGGACATACCCCT 2064
Qy 677 AspAlaLeuLysGluValValGlnLeuGluLysGlnLeuGlyGlyGlnLeuIleGlnGlu 696
Db 2065 TGTGCATTTTCAAGAGTGGTTTCAGATCAAGAGCATCAAGGTGGACAGCTCTCGAAGAA 2124
Qy 697 ProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspVal 716
Db 2125 CCAAAATGTGCAATTTCAAGAGGATACCTTTAGTCTTCAGATTTCTGCTCTTGATATT 2184
Qy 717 ProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHis 736
Db 2185 CCCCATTCTCTGGAGATTAACCAATTCAGTCCCTGCGAGGAGTCCCGTCTTCCCGC 2244
Qy 737 IleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerPro 756
Db 2245 GTGTGTGCAGTAAACCGGAGCCCTGCACTGTGCTTCTCCCTGGAGCGTTATACGCC 2304
Qy 757 SerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSer 776
Db 2305 ACTACCAACCCAGCTGCTGCTGCAAAATCTGCAATTCGGCAGCTCAAGGCCATGAACATC 2364
Qy 777 PheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGlu 796
Db 2365 CTCCAAGTGCAGATCAATCTTAGAGAGTGAACGAAACCATCATCTTTCTTCGCACAA 2424
Qy 797 SerGluAlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheLeuIle 816
Db 2425 GAGGACAGCATTTTCCTGTCGACAGACTGGCCCCAAAGCCCTTCAAAATTCCTACTCCATC 2484
Qy 817 ArgGlnLysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThr 836
Db 2485 AGACAGCGGATTTGTCTCATTTGATACCCCAATGCCAAAGGCCAAGGACTGGCAGATG 2544
Qy 837 LeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLysProSerPro 856
Db 2545 TTAGCACAGAAACACATCAACAGGATTTATCTTATTTGCTACAAAGATAGCCCA 2604
Qy 857 ThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsnLeuSerGln 876
Db 2605 TCTGTGTCTATTTGAACCTGTGGAGCTCGTCATCAGCATGATGTTGATCTTGACTCC 2664
Qy 877 LeuAlaAlaAlaAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThrValSerGlu 896
Db 2665 CTGGCTGTGCTTTGAAGAGATTGGAGGACACACACAGAAACTCTCAAAACATTTTCAGAA 2724
Qy 897 AlaGlu 898
Db 2725 TCCCAG 2730

RESULT 10
AY411748
LOCUS
DEFINITION
Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY411748
VERSION
AY411748.1 GI:39767716
```

```
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes (chimpanzee)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 2532)
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 2532)
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..2532
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1_>2532
/locus_tag="HCM4327"
gene
ORIGIN
Alignment Scores:
Pred. No.: 3,76e-157 Length: 2532
Score: 1890.00 Matches: 390
Percent Similarity: 60.87% Conservative: 125
Best Local Similarity: 46.10% Mismatches: 275
Query Match: 39.48% Indels: 56
DB: 29 Gaps: 16
US-09-970-944-2 (1-899) x AY411748 (1-2532)
```

```
Qy 103 GluValArgIleAsnValSerArgGlnValGluLysValPheGlyLeuGluGluTyr 122
Db 13 GAGGTGCAGATCGAGGTGTGGCGCAGCAGAGTGGAGAGCTCTTTGGCTGGAGGATTAC 72
Qy 123 TrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142
Db 73 TGGTGCAGTGTGGCTGGAGCTCTGGGCGACCCACCAAGAGTCCGCGAGCTACGTC 132
Qy 143 ArgIleAlaArgLeuArgLysAsnPheGluGlnGluProLeuAlaLysGluValSerLeu 162
Db 133 CGCATCGCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 192
Qy 163 GluGlnGlyIleValLeuProCysArgProProGluGlyIleProProAlaGluValGlu 182
Db 193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 252
Qy 183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202
Db 253 TGGCTCAAGATGAGGATGTCTATCGACCCCAACCCAGGACACCAACTTCCTGCTACCATC 312
Qy 203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222
Db 313 GACCACACCTCATCATCGCGAGCGCCCTGTGCGACACTGCCAATCTATACCTGCGGTG 372
Qy 223 AlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsn 242
Db 373 GCCAAGAACATCGTGGCCAAACGCCGAGGACCACTGCCACCGTCATCGTCTACGTGAAT 432
Qy 243 GlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyArgGlyTrp 262
```

Qy	576	MetAspHisCysGlyGluProSerProAspSerTrpSerLeuArgLeuLysGlnSer	595
Dd	1558	ATGCCCACTGTGCCGAAGTCAGTCCCGTACATGGATCTTTAGCTCAAGACCAGGCC	1617
Qy	596	CysGlySerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeu	615
Dd	1618	CACCAGGCCACTGGGAG---GAGGTGTACCTGGATGAGGAGACCTTGAACACACCC	1674
Qy	616	TyrTyrCysGlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuArgPhe	635
Dd	1675	TGCTACTGCCAGCTGGAGCCAGGCGCTGTACATCTCTGTGTGACAGCTGGGACCTAC	1734
Qy	636	AlaLeuValGlyGluAlaLeuSerValAlaAlaLysArgLeuLysLeuLeuPhe	655
Dd	1735	GTGTTTCAGGCGAGTCTTATTCNNNNNNNAGTCNNNNNGCTCCNGCTGGCGTCTTN	1794
Qy	656	AlaProValAlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThr	675
Dd	1795	GCCCTCCCTCTGCACCTCCCTGANNNNNGCCTCCGGGTACTACTGCCTGGNNNNNN	1854
Qy	676	HisAspAlaLeuLysGluValGlnLeuGlnLysGlnLeuGlyGlnLeuLeuGln	695
Dd	1855	NNNNNNNNNNNNAGGTGCTGGAGTCGAGCGGACTCTGGCGGATATCTTGGTGAG	1914
Qy	696	GluProArgValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAsp	715
Dd	1915	GAGCGAAACCGCTAAATGTTCAAGACAGTTTACCACCACTGGCGCTCTCCCTCATGAC	1974
Qy	716	ValProSerSerLeuTrpLysSerLysLeuLeuValSerTyrGlnGlnLeuProPheTyr	735
Dd	1975	CTCCCCATGCCATTTGGAGGAGCAAGCTGTGGCCAAATACCAGGAGATCCCTCTCTGT	2034
Qy	736	HisIleTrpAsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSer	755
Dd	2035	CACATTTGGAGTGGCAGCAGAGCCCTCCACTGCATTTTCACCTGGAGGAGGCACAGC	2094
Qy	756	ProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGln	775
Dd	2095	TTGGCTCCACAGACTCACCTGCAAGATCTGCGTGGCGCAAGTGGAGGGGAGGCCAG	2154
Qy	776	SerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuAlaLeu	795
Dd	2155	ATATTCCAGCTGCATACCACCTGTGCA---GAGACACCTGTGGCTCCCTGGACACTCTC	2211
Qy	796	GluSerGluAlaGly-----ValProAlaLeuValGlyProSerAlaPheLysIlePro	813
Dd	2212	TGCTCTGCCCTGGCAGCAGCTGTACACCCACCTGGGACCTTATGCTCTCAAGATCCCA	2271
Qy	814	PheLeuIleArgGlnLysIleSerSerLeuAspProProCysArgArgGlyAlaAsp	833
Dd	2272	CTGTCCATCCGCGAAGATATGCACAGCTAGATGCCCCCACTCACGGGCAATGAC	2331
Qy	834	TrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePheAlaSerLys	853
Dd	2332	TGGCGATCTTAGCAGACAGCTCTCTATGGACCGGTACTGAATTACTTTGCCACCAA	2391
Qy	854	ProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPheProAsnGlyAsn	873
Dd	2392	GCGAGCCCAACGGTGTGATCTCTGGACCTCTGGAGAGTCTGCAGCAGGACGATGGGAC	2451
Qy	874	LeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGlyLeuPheThr	893
Dd	2452	CTCAACAGCTGGCGAGTGCCTTGGAGGAGTGGGCAAGATGAGATGCTGTGTGGTG	2511
Qy	894	ValSerGluAlaGluCys	899
Dd	2512	GCCACCGAGGGGACTGC	2529
RESULT 11			
BC033727		1532 bp	mRNA
LOCUS			linear
DEFINITION			Homo sapiens netrin receptor Unc5h1, mRNA (cdna clone
			HTC 19-NOV-2003



QY 197 nValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAl 217  
Db 911 TGTATACATCAGCGGAGCACAGCCTGGTGTGCGACAGCGCCGCTTGTGACACGGC 970  
QY 217 aAsnTyrThrCysValAlaLysAsnIleValAlaArgArgArgSerAlaSerAlaVala 237  
Db 971 CAACACTACCTGGTGGCCAAAGAACATCGTGGCACGTGGCGGAGCGCTCCCTGCTGT 1030  
QY 237 lIleValTyrValAsnGlyGlyTyrSerThrTyrThrGluTyrSerValCysSerAlaSe 257  
Db 1031 CATCGTCTAC----- 1040  
QY 257 rCysGlyArgGlyTyrPdnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGl 277  
Db 1040 ----- 1040  
QY 277 yGlyAlaPheCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSe 297  
Db 1040 ----- 1040  
QY 297 rValAspGlySerTyrPdnLysArgSerArgSerCysThrAlaCysGlyLeuAspCysThrHi 317  
Db 1041 -GTGACCGGAGCTGGAGCCGCTGGAGAAAGTGGTGGCGCTGTGGGTGACCTGCACCCA 1099  
QY 317 sTyrArgSerArgGluCysSerAspProAlaProArgAsnGlyGlyGluGluCysGlnGl 337  
Db 1100 CTGGCGGAGCGGTGAGTGTCTGTACCCAGCACCCCGCAAGAGGAGGAGGAGTCCCGAGG 1159  
QY 337 yThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHisSerAlaSerGlyPr 357  
Db 1160 CACTGACCTGGACACCCGCACTGTACAGTGTCTGTGTGTACACTGTCTTGTGGCCC 1219  
QY 357 oGluAspValAlaLeuTyrValGlyLeuIleAlaValAlaValCysLeuValLeuLeuLe 377  
Db 1220 TGAGGAGTGGCCCTCTATGTGGCCCTCATCGCGCTGCTGCTGCTGCTGCTGCTGCT 1279  
QY 377 uLeuValLeuLeuValTyrCysArgLysGluGlyLeuAspSerAspValAlaAs 397  
Db 1280 GCTTGTCTCATCTCTGTTATTGCGGAAGAAGAGGAGGCTGGACTCAGATGTGGCTGA 1339  
QY 397 pSerSerIleLeuThrSerGlyPheGlnProValSerIleLysProSerLysAlaAspAs 417  
Db 1340 CTCGTCACTTCTACCTCAGGCTTCCAGCCCGTTCAGATCAAGCCAGCAAGACAGCAA 1399  
QY 417 nProHisLeuLeuThrIleGlnProAspLeuSer---ThrThrThrThrThrGlnGlySe 436  
Db 1400 CCCCCATCTGCTCACCATCCAGCGGACCTCAGCACCCACCACCTACCTACCGGGCAG 1459  
QY 436 rLeuCysProArgGlnAspGlyProSerProLysPheGlnLeuThrAsnGlyHisLeuLe 456  
Db 1460 TCTCTGTCGGGAGGATGGGCCAGGCCCAAGTTCAGGCTCACCATTGGGCACCTGCT 1519  
QY 456 uSerProLeuGly 460  
Db 1520 CAGCCCCCTGGGT 1532  
RESULT 12  
AY401470 2507 bp DNA linear GSS 12-DEC-2003  
LOCUS AY401470  
DEFINITION Pan troglodytes HCM0901 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY401470  
VERSION AY401470.1 GI:39757459  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE 1 (bases 1 to 2507)  
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 2507)  
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source 1..2507  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
gene <1..>2507  
/locus\_tag="HCM0901"  
ORIGIN  
Alignment Scores: 9.18e-127 Length: 2507  
Pred. No.: 1550.00 Matches: 332  
Score: 52.59% Conservative: 135  
Percent Similarity: 37.39% Mismatches: 257  
Best Local Similarity: 32.38% Indels: 164  
Query Match: 29 Gaps: 14  
DB: 14  
US-09-970-944-2 (1-899) x AY401470 (1-2507)  
QY 103 GluValArgIleAsnValSerArgGlnGlnValGlnLysValPheGlyLeuGluGluTyr 122  
Db 15 GAAGTGTTCATCATATGTTACAGGCAACAGGTGGAGGACTTCCATGGGCCCGCAGACTAT 74  
QY 123 TrpCysGlnCysValAlaTyrSerSerSerGlyThrThrLysSerGlnLysAlaTyrIle 142  
Db 75 TGGTGGCCAGTGTGTGGGTGGAGCCACTGGGTACTTCCAGAGCAGGAGGAGGCTCTGTG 134  
QY 143 ArgIleAlaArgLeuArgLysAsnPheGlnGlnGluProLeuAlaLysGluValSerLeu 162  
Db 135 CGCATAGCTATTATTACGMAAACTTTGAAACAGACCACCAAGGAAGGAGGTTCCCAT 194  
QY 163 GluGlnGlyIleValLeuProCysArgProGluGlyIleProAlaGluValGlu 182  
Db 195 GAAGGCATGATTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 254  
QY 183 TrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIleThrArg 202  
Db 255 TGGCTGAAAATGAAGACCCATTCATCTGACACAGACGAGAACATTGACACAGGGCT 314  
QY 203 GluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThrCysVal 222  
Db 315 GACCAATACTGATCATCAGCGAGGACGCTCTCGGACTCAGGAAATATACACCTGCATG 374  
QY 223 AlaLysAsnIleValAlaArgArgSerAlaSerAlaAlaValIleValTyrValAsn 242  
Db 375 GCAGCAACATCGTGGCTTAAGAGGAGAGAGCCCTGTCGGCCACTGTTGTGTCTACGTGAAT 434  
QY 243 GlyGlyTyrSerThrTyrThrGluTyrSerValCysSerAlaSerCysGlyArgGlyTyr 262  
Db 435 GGAGGCTGGTCTTCTCTGGACAGAGTGGTTCAGCTGCAATGTTGCTGTGTAGAGGATGG 494  
QY 263 GlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPheCysGlu 282  
Db 495 CAGAAACGCTCCGAGACCTGCACCAACCCAGCTCCTCAATGGTGGGCGCTTTGTGTAG 554  
QY 283 GlyGlnAsnValHisAspArgThrValSerSerLeuLeuValSerValAspGlySerTrp 302  
Db 555 GAAATGTCTAGTCAGAAATAATACC----- 578

```
QY 303 SerProTrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgSerArgGlu 322
Db 578 -----
QY 323 CysSerAspProAlaProArgAsnGlyGlyGluCysGlnGlyThrAspLeuAspThr 342
Db 578 -----
QY 343 ArgAsnCysThrSerAspLeuCys----- 350
Db 579 -----TGCACCTCT-----CITTGCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 629
QY 350 ----- 350
Db 630 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 689
QY 350 ----- 350
Db 690 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 749
QY 351 -----ValHisSerAlaSerGlyProGluAspValAlaLeuTyValGly 365
Db 750 NNNNNNNNNNNNGCATTTGAGATGCCAGC-----GACATTCCTTTGTACTCGGC 800
QY 366 LeuLeuAlaValAlaValCysLeuValLeuLeuValLeuValLeuValLeuValTyCys 385
Db 801 TTG---GGCGCTCGCGTCGTCGCGCTGTCAGTCCTGGTCATTGGTGTACCCCTTTACAG 857
QY 386 ArgLysGlyGluGlyLeuAspSerAspValAlaAspSerSerIleLeuThrSerGlyPhe 405
Db 858 CGGAGCCAGAGTACTATGCGTGGAGCTCATTTGACTCTTCTGCAATTGACAGGTGGCTTC 917
QY 406 GlnProValSerIleLysProSerLysAlaAspAsnProHisLeuLeu-----ThrIle 423
Db 918 CAGACCTTCAACTTCAAAACAGTCCGCTCAAGNNNACTCCCTGCTGCTGAATTCGGCCATG 977
QY 424 GlnProAspLeuSerThrThrThrThrTyrglnGlySerLeuCysProArgGlnAspGly 443
Db 978 CAGCCAGATCTGACAGTGGCCGGACATACAGCGGACCCATCTGT---CTGCGAGGAC--- 1031
QY 444 ProSerProLysPheGlnLeuThrAsnGlyHisLeuLeuSerProLeuGly----- 460
Db 1032 CCTCTGGACAGAGAGCTCATGACAGAGTCTCTACTGTTTAAACCTTTGTGCGACATCAMA 1091
QY 460 ----- 460
Db 1092 GTCAAAAGTCCAGAGCTGTTTCATGGTTTCCCTGGGAGTGTCTGAGAGAGCTGAGTACCAC 1151
QY 461 -----GlyGlyArgHisThrLeuHisHisSerSer 470
Db 1152 GGCAAGATCATTTCCAGGACTTTTCCCATGGAAACAACACAGCTTTAGTACAATGCAT 1211
QY 471 ProThrSerGluAlaGluGluPheValSerArgLeuSerThrGlnAsnTyrrPheArgSer 490
Db 1212 CCCAGAAATAAATG---CCCTACATCCAAATCTGTCA-----TCA 1250
QY 491 LeuProArgGlyThrSerAsnMetThrTyrglyThrPheAsnPheLeuGlyGlyArgLeu 510
Db 1251 CTCCCCACAGGACAGAACTGAGGACAACTGTGTGTCTNNNGCCATTTAGGGGGCGGNNNN 1310
QY 511 MetIleProAsnThrGlyIleSerLeuLeuIleProProAspAlaIleProArgGlyLys 530
Db 1311 NNNNNN---NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1367
QY 531 IleTyrgluIleTyrrLeuThrLeuHisLysProGluAspValArgLeuProLeuAlaGly 550
Db 1368 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1427
QY 551 CysGlnThrLeuLeuSerProIleValSerCysGlyProGlyValLeuThrArg 570
Db 1428 TCTGAGGTGCTCTCTGAGTCTCTGAAGTCACTGTGTGTCTCTCCAGACATGATCGTCACCAT 1487
```

```
QY 571 ProValIleLeuAlaMetAspHisCysGlyGluProSerProAspSerTrpSerLeuArg 590
Db 1488 CCCTTTGCATTGACATCCCCACACTGTGCAGATGTCTAGTNNNGAGCACTGGAAATATCAT 1547
QY 591 LeuLysLysGlnSerCysGlyGlySerTrpGluGlnAspValLeuHisLeuGlyGluGlu 610
Db 1548 TTAAGAAGAGGACACACAGCAGGCAATGGAG---GAAGTGAATGTGAGTGAAGATGAA 1604
QY 611 AlaProSerHisLeuTyrrCysGlnLeuGluAlaSerAlaCysTyrrValPheThrGlu 630
Db 1605 TCTACATCC-----TGTTACTGCTTTTGGACCCCTTTGCACTGTCTGCTCTGGAC 1658
QY 631 GlnLeuGlyArgPheAlaLeuValGlyGluAlaLeuSerValAlaAlaLysArgLeu 650
Db 1659 AGCTTTGGGACCTATCGCTCACTGGAGAGCCCAATCACAGACTGTGCGGTGAAGCAACTG 1718
QY 651 LysLeuLeuLeuPheAlaProValAlaCysThrSerLeuGluTyrrAsnIleArgValTy 670
Db 1719 AAGGTGGCGGTTTTTGGCTGCATGTCTGTAACTCCCTGGATTACAACTTGAGAGTTTAC 1778
QY 671 CysLeuHisAspThrHisAspAlaLeuLysGluValValGlnLeuGlyLysGlnLeuGly 690
Db 1779 TGTGTGGACAATACCCCTTGTGCATTTCCANGAAGTGGTTTCAGATGAAAGGCATCAAGGT 1838
QY 691 GlyGlnLeuIleGlnGluProArgValLeuHisPheLysAspSerTyrrHisAsnLeuArg 710
Db 1839 GGACAGCTCTGTGAAGAACCAAAATTTGCTGCATTTCCAAAGGAATACCTTTAGTCTTCAG 1898
QY 711 LeuSerIleHisAspValProSerSerLeuTyrrLysSerLysLeuLeuValSerTyrrGln 730
Db 1899 ATTTCGTCTCTGTATTTCCCCATTTCTCTGGAGAAATTAACCATTCCTACTGCTGCCAG 1958
QY 731 GluLeuProPheTyrrHisIleTrpAsnGlyThrGlnArgTyrrLeuHisCysThrPheThr 750
Db 1959 GAAGTCCCGTCTCTCCCGCTGTGGTGCAGTAAACCGGACGCCCTGCTGCTGCTCTTCTCC 2018
QY 751 LeuGluArgValSerProSerThrSerAspLeuAlaCysLysLeuTrpValTrpGlnVal 770
Db 2019 CTGGAGCGTATACGCCCACTACCCAGCTGTCTCTGCAAAATCTGCATTCGGCAGCTC 2078
QY 771 GluGlyAspGlyGlnSerPheSerIleAsnPheAsnIleThrLysAspThrArgPheAla 790
Db 2079 AAAGGCCATGAACAGATCTCTCAAGTGCAGACATCAATCTAGAGAGTGAACGAGAAACC 2138
QY 791 GluLeuLeuAlaLeuGluSerGluAlaGlyValProAlaLeuValGlyProSerAlaPhe 810
Db 2139 ATCACTTTCTTTGCAAGAGGACAGCACTTTCCCTGTCACAGACTGGCCCCAAAGCTTC 2198
QY 811 LysIleProPheLeuIleArgGlnLysIleIleSerSerLeuAspProProCysArgArg 830
Db 2199 AAAATTCCTACTCTCCATCAGACGCGGATTTGTGTACATTTGATACCCCAATGCCAAA 2258
QY 831 GlyAlaAspTrpArgThrLeuAlaGlnLysLeuHisLeuAspSerHisLeuSerPhePhe 850
Db 2259 GCGAAGACTGCGACATGTTAGCACAGAAAACACGATCAACAGGAATTTATCTATTTC 2318
QY 851 AlaSerLysProSerProThrAlaMetIleLeuAsnLeuTrpGluAlaArgHisPhePro 870
Db 2319 GCTACACAAAGTAGCCCATCTGCTCATTTTGAACCTGTGGAGAGTCTGTCATCAGCAT 2378
QY 871 AsnGlyAsnLeuSerGlnLeuAlaAlaValAlaGlyLeuGlyGlnProAspAlaGly 890
Db 2379 GATGTGTATCTTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2438
QY 891 LeuPheThrValSerGluAlaGlu 898
Db 2439 CTCTCAAACTTTCAGATCCAG 2462
```

RESULT 13  
BI758231  
LOCUS  
DEFINITION

603029876F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5200171 5',  
mRNA sequence.

1034 bp mRNA linear

EST 25-SEP-2001



ACCESSION BI758231  
 VERSION BI758231.1 GI:15749809  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1034)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabps-remail.nih.gov](mailto:cgabps-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA library preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM11501 row: g column: 20  
 High quality sequence stop: 793.  
 Location/Qualifiers  
 1..1034  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5200171"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_114"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC Library."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,13e-111 Length: 1034  
 Score: 1373.50 Matches: 291  
 Percent Similarity: 88.82% Conservative: 11  
 Best Local Similarity: 85.59% Mismatches: 28  
 Query Match: 28.69% Indels: 14  
 DB: 12 Gaps: 2  
 US-09-970-944-2 (1-899) x BI758231 (1-1034)  
 QY 36 ProGlyAlaAsnProAspLeuValGluProHisPheLeuValGluProGluAspValTyrIle 55  
 Db 2 CCTGGTCCACCCGGACCTCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACATC 61  
 QY 56 ValLysAsnLysProValLeuValCysLysAlaValProAlaThrGlnIlePhePhe 75  
 Db 62 GTCAAGAACCAAGCCAGTGTCTTGTGTGCAAGCCGTCGCCGCCAGCATCTTCTTC 121  
 QY 76 LysCysAsnGlyGluTTPValArgGlnValAspHisValIleGluArgSerThrAspGly 95  
 Db 122 AAGTGCACCGGGAGTGGTGGCCAGGTGGACCAACGATCATGATCGAGCGAGCACAGCGG 181  
 QY 96 SerSerGlyGluProThrMetGluValArgIleAsnValSerArgGlnValGluLys 115  
 Db 182 AGCAGTGGGCTGCCACCATGGAGTCCGCATTATGTCTCAAGGCAGCAGGTGAGAG 241  
 QY 116 ValPheGlyLeuGluGluTyrTrpCysGlnCysValAlaTTPSerSerGlyThrThr 135  
 Db 242 GTGTTCGGGCTGGAGGAATATGTGGTCCAGTGGGATGAGTCTCTCGGGACCAACC 301  
 QY 136 LysSerGlnLysAlaTyrIleArgIleAlaArgLeuArgLysAsnPheGluGlnGluPro 155

Db 302 AAGAGTCAGAGGCCCTACATCCGCATAGCTATTTTCGCCAGAGAACTTCGAGCAGGAGCCG 361  
 QY 156 LeuAlaLysGluValSerLeuGlnGlyIleValLeuProCysArgProGluGly 175  
 Db 362 CTGGCCAGGAGGTGCTCTGAGCAGGACATCGTGTCCCTCCCTCCGCTCACCGAGGCG 421  
 QY 176 IleProProAlaGluValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAsp 195  
 Db 422 ATCCCTCCAGCCGAGGTGGAGTGGCTCCGAAACGAGGACCTGTGTGACCCGTCCTGGAC 481  
 QY 196 ProAsnValTyrIleThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAsp 215  
 Db 482 CCCAATGTATACATCAGCGGGAGCACAGCTGTGTGTCGACAGCCGCTTGTGAC 541  
 QY 216 ThrAlaAsnTyrThrCysValAlaLysAsnIleValAlaArgArgSerAlaSerAla 235  
 Db 542 ACGGACAACATACACCTGCGTGGCCAGAAACATCGTGGACGCTGCCGAGCGCTCGCT 601  
 QY 236 AlaValIleValTyrValAsnGlyGlyTrpSerThr-TripThrGluTrpSerValCysSe 255  
 Db 602 GCTGTCACTGTCTACGTGAACGGTGGTGTGACCGTGGACCGAGTGGTCCGCTTCGCG 661  
 QY 255 rAlaSerCysGlyArgGlyTrpGlnLys-ArgSerArgSerCysThrAsnProAlaProL 275  
 Db 662 -GCCAGCTGTGGCGCGGCTGGCAGAAACCGGAGCGGAGTGCAC-AAACCCGCTGCCTC 719  
 QY 275 euAsnGlyGlyAlaPheCysGlu-GlyGlnAsnValHisAspArgThrValSerSerLeu 294  
 Db 720 TCAACGGGGCGCTTCTGTGAGGGGCGAGATGTC-----CAGAAAGCAGTGGCGCCA 773  
 QY 295 LeuValSerValAspGlySerTrpSerProTrpSerLysTrpSerAlaCys--GlyLeuA 314  
 Db 774 CTTGTGCCAGTGAGCGGACGCTAGCACGTGGAGCCAGTGGTGGCTGTGGGCTTGG 833  
 QY 314 spCysThrHis-TrpArgSerArgGluCysSerAsp-ProAlaProArgAsnGlyGlyG 333  
 Db 834 CTTTGACCCCACTTGGCGCGGAGTGTCTCTGAACCCAGACCCCGGACCGGACAGGGG 893  
 QY 333 uGluCys--GlnGlyThrAspLeuAspThrArgAsnCysThrSerAspLeuCysValHi 352  
 Db 894 GAGGTGTGCCGCTCCTGACCTGGACCCCGGGATGG-TCCAGTGAGTCTGTGTCCC 952  
 QY 352 sSerAlaSerGlyProGluAspValAlaLeuTyrValGlyLeuIle 367  
 Db 953 -ACGGTTCCTGGCCCTGAGGACTTGGCTCTCTAATGTGGGCTCATC 997  
 RESULT 14  
 BI818609  
 LOCUS 818 bp mRNA linear EST 04-OCT-2001  
 DEFINITION 60303362F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5174559 5', mRNA sequence.  
 ACCSSION BI818609  
 VERSION BI818609.1 GI:15929902  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 818)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabps-remail.nih.gov](mailto:cgabps-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLAM11434 row: 1 column: 16  
 High quality sequence stop: 744.



FEATURES  
source

Location/Qualifiers  
1. .818  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5174559"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_115"  
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 5,39e-109 Length: 818  
Score: 1341.50 Matches: 263  
Percent Similarity: 80.61% Conservative: 3  
Best Local Similarity: 79.70% Mismatches: 5  
Query Match: 28.02% Indels: 59  
DB: 12 Gaps: 1

US-09-970-944-2 (1-899) x B1818609 (1-818)

QY 1 MetAlaValArgProGlyLeuTrpProAlaLeuGlyLeuValLeuAlaAlaTriLeu 20  
DB 4 ATGGCGCTCCGGCCCGGCTGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCTGCTC 63  
QY 21 ArgGlySerGlyAlaGlnGlnSerAlaThrValAlaAsnProValProGlyAlaAsnPro 40  
DB 64 CGCGGCTCGGGTGGCCAGAGAGTGCCACCGTGGCCCAACCGAGTGGTGGCCACCGC 123  
QY 41 AspLeuLeuProHisPheLeuValGluProGluAspValTyrLeuValValAsnLysPro 60  
DB 124 GACCTGCTCCCGACTCTCTGGTGGAGCCCGAGATGTACATCGTCAAGAACAGCCCA 183  
QY 61 ValLeuLeuValCysLysAlaValProAlaThrGlnIlePhePheLysCysAsnGlyGlu 80  
DB 184 GTGCTGCTGTGTGCAAGCGGTGGCCGCGCCAGCAGATCTTCAAGTGCACAGCGGGAG 243  
QY 81 TrpValArgGlnValAspHisValIleGluArgSerThrAspGlySerGlyGluPro 100  
DB 244 TGGTGGCGCGAGTGGACACGATGATGCGAGCGAGCAGCAGCGGAGCAGTGGGCTGCC 303  
QY 101 ThrMetGluValArgIleAsnValSerArgGlnGlnValGluLysValPheGlyLeuGlu 120  
DB 304 ACCATGGAGTCCGCATTATGTCTCAAGCAGCAGGTCCAGAGGTGTTCGGGCTGGAG 363  
QY 121 GluTyrTrpCysGlnCysValAlaTrpSerSerGlyThrThrLysSerGlnLysAla 140  
DB 364 GAATACTGTGCGCAGTGGCATGGAGTCTCTCGGCGCACCAACAGAGTCAAGAGGCC 423  
QY 141 TyrIleArgIleAlaArgLeuArgLysAsnPhelGluGlnProLeuAlaLysGluVal 160  
DB 424 TACATCCGATAGCTATTATTGGCAAGAACTTCGAGCAGAGCGCGCTGGCCAGAGAGTG 483  
QY 161 SerLeuGluGlnGlyLeuValLeuProCysArgProProGluGlyLeuProAlaGlu 180  
DB 484 TCCCTGGAGCAGGCGATGTGTCTGCTCCCTGCCGTCACCGAGGCGATCCCTCCAGCCGAG 543  
QY 181 ValGluTrpLeuArgAsnGluAspLeuValAspProSerLeuAspProAsnValTyrIle 200  
DB 544 GTGGAGTGGCTCCGGAACGAGGACCTGTGGACCGCTCCCTGGACCCCAATGTATACATC 603  
QY 201 ThrArgGluHisSerLeuValValArgGlnAlaArgLeuAlaAspThrAlaAsnTyrThr 220  
DB 604 ACGCGGAGCAGCGCTGTGTGGCAGAGCGCGCTTGTCTGACACGCGCAACTACACC 663

QY 221 CysValAlaLysAsnIleValAlaArgArg-ArgSerAlaSerAlaValIleVal 240  
DB 664 TGCCTGGCCAAACATCGTGCACGTCAGCCGAGGCGCTCCGCTGCTGCTGCTGCTA 723  
QY 240 rValAsnGlyGlyTrpSerThrTrpThrGluTrpSerValCysSerAlaSerCysGlyAr 260  
DB 724 C----- 724  
QY 260 gGlyTrpGlnLysArgSerArgSerCysThrAsnProAlaProLeuAsnGlyGlyAlaPh 280  
DB 724 ----- 724  
QY 280 eCysGluGlyGlnAsnValHisAspArgThrValSerSerLeuValSerValAspGly 300  
DB 725 -----GTGGACGG 732  
QY 300 YSerTrpSerPro-TrpSerLysTrpSerAlaCysGlyLeuAspCysThrHisTrpArgS 320  
DB 733 CAGCTGGAGCCGTTGGAGCCAGTGGTGGCTGTGGCTGTGGCTGACCCACTGGCGGA 792  
QY 320 erArgGluCysSerAspProAlaPro 328  
DB 793 GCCGGAAGTGTCTGACCCCAACCCG 818

## RESULT 15

## BQ689148

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

BQ689148 874 bp mRNA linear EST 15-JUL-2002  
AGENCOURT 8064396 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6206234  
5', mRNA sequence.

BQ689148

BQ689148

BQ689148.1

GI:21814464

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 874)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2360 row: 9 column: 03

High quality sequence stop: 737.

Location/Qualifiers

1. .874

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6206234"

/tissue\_type="ductal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_110"

/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dr priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

## Pred. No.:

## Length:

## 874

Score: 1274.00 Matches: 255  
Percent Similarity: 63.59% Conservative: 0  
Best Local Similarity: 63.59% Mismatches: 0  
Query Match: 26.61% Indels: 146  
DB: 13 Gaps: 1

US-09-970-944-2 (1-899) x BQ689148 (1-874)  
QY 499 ThrTyrGlyThrPheAsnLeuGlyGlyArgLeuMetIleProAsnThrGlyIleSer 518  
Db 1 ACCTATGGGACCTTCACCTTCCTCGGGGCGGCTGATGATCCCTAATACAGGAATCAGC 60  
QY 519 LeuLeuIleProProAspAlaIleProArgGlyIleTyrGluIleTyrLeuThrLeu 538  
Db 61 CTCCTCATCCCCCAGATGCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCGTG 120  
QY 539 HisGlyProGluAspValArgLeuProLeuAlaGlyCysGlnThrLeuLeuSerProIle 558  
Db 121 CACAAGCCGGAAGACGTGAGGTGCCCTAGCTGGCTGCAGACC----- 165  
QY 559 ValSerCysGlyProProGlyValLeuLeuThrArgProValIleLeuAlaMetAspHis 578  
Db 165 ----- 165  
QY 579 CysGlyGluProSerProAspSerTrpSerLeuArgLeuLysLysGlnSerCysGluGly 598  
Db 165 ----- 165  
QY 599 SerTrpGluGlnAspValLeuHisLeuGlyGluGluAlaProSerHisLeuTyrTyrCys 618  
Db 165 ----- 165  
QY 619 GlnLeuGluAlaSerAlaCysTyrValPheThrGluGlnLeuGlyArgPheAlaLeuVal 638  
Db 165 ----- 165  
QY 639 GlyGluAlaLeuSerValAlaAlaAlaLysArgLeuLysLeuLeuPheAlaProVal 658  
Db 165 ----- 165  
QY 659 AlaCysThrSerLeuGluTyrAsnIleArgValTyrCysLeuHisAspThrHisAspAla 678  
Db 165 ----- 165  
QY 679 LeuLysGluValValGlnLeuGluLysGlnLeuGlyGlnLeuIleGlnProArg 698  
Db 165 ----- 165  
QY 699 ValLeuHisPheLysAspSerTyrHisAsnLeuArgLeuSerIleHisAspValProSer 718  
Db 166 ---CTGCACCTTCAAGGACAGTTACCAACCTGGCGCTTATCCATCCAGATGGCCAGC 222  
QY 719 SerLeuTrpLysSerLysLeuLeuValSerTyrGlnGluIleProPheTyrHisIleTrp 738  
Db 223 TCCCTGTGGAAGAGTAAGTCTCTGTGACGTACCAAGGAGATCCCTTTATCAGATCTGG 282  
QY 739 AsnGlyThrGlnArgTyrLeuHisCysThrPheThrLeuGluArgValSerProSerThr 758  
Db 283 AATGGCAGCGCGGTACTTGCACCTGCACCTTACCCCTGGAGCGTGTCCAGCCCCAGCACT 342  
QY 759 SerAspLeuAlaCysLysLeuTrpValTrpGlnValGluGlyAspGlyGlnSerPheSer 778  
Db 343 AGTGACCTGGGCTGCAAGCTGTGGGTGGGAGGTGGAGGGGCGGCGAGAGCTTCAGC 402  
QY 779 IleAsnPheAsnIleThrLysAspThrArgPheAlaGluLeuLeuAlaLeuGluSerGlu 798  
Db 403 ATCAACTTCAACATCACCAAGGACACCAAGTTTGTGAGCTGCTGGCTCTGGAGAGTAA 462  
QY 799 AlaGlyValProAlaLeuValGlyProSerAlaPheLysIleProPheIleuIleArgGln 818  
Db 463 GCGGGGGTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCCTCATTCGGGAG 522  
QY 819 LysIleIleSerSerLeuAspProProCysArgArgGlyAlaAspTrpArgThrLeuAla 838

Search completed: October 6, 2004, 22:37:02  
Job time : 5582 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 09:39:46 ; Search time 11136 seconds  
(without alignments)  
11213.297 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccgggtgag.....ccttccccacacggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2881	100.0	2881	6	AX527916 Sequence
2	2676.4	92.9	2752	6	AX449572 Sequence
3	2615	90.8	2697	6	AX451652 Sequence
4	2402.6	83.4	3580	6	AX367094 Sequence
5	2327.8	80.8	3992	10	MMU487852
6	2235.8	77.6	3014	6	BD057524
7	2189.4	76.0	2697	6	AX268596 Sequence
8	2189.4	76.0	2697	10	RNU87305
9	2020.6	70.1	3844	10	BC058084
10	1691.4	58.7	2688	9	BC009333
11	1570.4	54.5	1787	6	BD057525
12	1291.2	44.8	9700	6	AX054976
13	940.2	32.6	9299	10	MMU72634
14	938.4	32.6	2962	5	AY187310
15	936.8	32.5	9328	10	AB118026
16	916	31.8	3646	9	AF055634
17	895	31.1	3770	9	AY126437
18	889.4	30.9	4294	10	AK122575
19	889	30.9	2860	6	AX686445
20	887.4	30.8	2860	6	AX686447 Sequence
21	875.8	30.4	2995	6	AX497288 Sequence
22	866.4	30.1	2895	6	AX512281
23	859.8	29.8	3672	10	BC057560
24	859.8	29.8	3672	10	BC048162
25	856.8	29.7	3933	9	AK128132
26	856.8	29.7	3935	9	AB096256
27	853.6	29.6	3884	6	AX464012 Sequence
28	853.6	29.6	3884	9	AY358351
29	821.6	28.5	3788	10	MMU487853
30	794.4	27.6	2838	10	RNU87306
31	783.4	27.2	2831	6	BD057526
32	735.8	25.5	813	6	AX054892 Sequence
33	665.6	23.1	2832	5	AY099459 Xenopus 1
34	597.6	20.7	2661	6	AX800717 Sequence
35	597.6	20.7	2868	6	AX800719 Sequence
36	573.4	19.9	604	6	AX451656 Sequence
37	560.6	19.5	3252	9	AY358147
38	557.8	19.4	7002	6	BD162193
39	557.8	19.4	7002	9	AB055056
40	553	19.2	3925	10	MMU487854
41	502.6	17.4	1438	6	AX054888
42	467.4	16.2	2406	6	AX833115
43	467.4	16.2	2406	9	AK094595
44	465.8	16.2	1948	6	AX882354 Sequence
45	465.8	16.2	1948	6	BD159676 Primer fo

# ALIGNMENTS

RESULT 1  
AX527916  
LOCUS AX527916 2881 bp DNA linear PAT 21-NOV-2002  
DEFINITION Sequence 1 from Patent WO0229038.  
ACCESSION AX527916  
VERSION AX527916.1 GI:25172359  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Herrmann,J.L., Rastelli,L. and Shimketa,R.A.  
TITLE Novel proteins and nucleic acids encoding same and antibodies directed against these proteins

JOURNAL	Patent: WO 0229038-A 1 11-APR-2002;
FEATURES	Curagen Corporation (US)
Location/Qualifiers	1..2881
source	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	100.0%; Score 2881; DB 6; Length 2881;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2881; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AGTGGGCTCGGGCTGAGGGCTAAAGCGCCCTCCCGCCCGCGGGGCGCCCGCCCG 60
DB	1 AGTGGGCTCGGGCTGAGGGCTAAAGCGCCCTCCCGCCCGCGGGGCGCCCGCCCG 60
QY	61 GCCGCCCGCTGCCCGCCCGGGCCATGGCCGTCCGGCCCGCCCTGTGGCCAGCGCTCC 120
DB	61 GCCGCCCGCTGCCCGCCCGGGCCATGGCCGTCCGGCCCGCCCTGTGGCCAGCGCTCC 120
QY	121 TGGGCATAGTCTCGCGCGCTTGCTCCGGCGCTCGGCTGCCAGCAGAGTGCCACCGTGG 180
DB	121 TGGGCATAGTCTCGCGCGCTTGCTCCGGCGCTCGGCTGCCAGCAGAGTGCCACCGTGG 180
QY	181 CCAACCCAGTGTGGTGCACACCCGGACCTGCTCCCGGCTCGGCTGCCAGCAGAGTGCCACCGTGG 240
DB	181 CCAACCCAGTGTGGTGCACACCCGGACCTGCTCCCGGCTCGGCTGCCAGCAGAGTGCCACCGTGG 240
QY	241 ATGTGTACATCGTCAAGAACCAAGCAGTGTGTTGTGTCAGAGGCGCTGCCCGCCACGC 300
DB	241 ATGTGTACATCGTCAAGAACCAAGCAGTGTGTTGTGTCAGAGGCGCTGCCCGCCACGC 300
QY	301 AGATCTTCTCAAGTGCACACGGGAGTGGTGGCCAGGTGGACCAAGTGTGATCGAGCGCA 360
DB	301 AGATCTTCTCAAGTGCACACGGGAGTGGTGGCCAGGTGGACCAAGTGTGATCGAGCGCA 360
QY	361 GCACAGCGGGAGCAGTGTGAGCCGACATGAGAGTCCCGATTAATGTCTCAAGGAGC 420
DB	361 GCACAGCGGGAGCAGTGTGAGCCGACATGAGAGTCCCGATTAATGTCTCAAGGAGC 420
QY	421 AGTCCAGAGAGTGTTCGGGCTGGAGGAATACTTGTGCCAGTGGCTGGCATGAGCTCCT 480
DB	421 AGTCCAGAGAGTGTTCGGGCTGGAGGAATACTTGTGCCAGTGGCTGGCATGAGCTCCT 480
QY	481 CGGGCACCAACAGAGTCAAGAGGCTTACATCCGATAGCCAGATTCGGCAAGAACTTCG 540
DB	481 CGGGCACCAACAGAGTCAAGAGGCTTACATCCGATAGCCAGATTCGGCAAGAACTTCG 540
QY	541 AGCAGAGCGCTGGCCAGAGAGTGTCCCTGGAGCAGGGCATCGTCTGCCCTGCCGTC 600
DB	541 AGCAGAGCGCTGGCCAGAGAGTGTCCCTGGAGCAGGGCATCGTCTGCCCTGCCGTC 600
QY	601 CACCGAGGGCATCCCTCCAGCCGAGGTGGCTCCGGAACGAGACCTGGTGACC 660
DB	601 CACCGAGGGCATCCCTCCAGCCGAGGTGGCTCCGGAACGAGACCTGGTGACC 660
QY	661 CGTCCCTGGAACCAATGTATACATCACGCGGAGCAGCGCTGGTGGCGACAGGCC 720
DB	661 CGTCCCTGGAACCAATGTATACATCACGCGGAGCAGCGCTGGTGGCGACAGGCC 720
QY	721 GCCTTGTGACAGCGGCAACTACACCTTGGTGGCCAGACATCGTGGCATCGTCCGCGCA 780
DB	721 GCCTTGTGACAGCGGCAACTACACCTTGGTGGCCAGACATCGTGGCATCGTCCGCGCA 780
QY	781 GCGCTCCGCTGTGTCATCTGTACCTGAACGCTGGTGGTTCGACCTGGACCGAGTGGT 840
DB	781 GCGCTCCGCTGTGTCATCTGTGTACCTGAACGCTGGTGGTTCGACCTGGACCGAGTGGT 840
QY	841 CCGTGTGAGCGCCAGTGTGGGCGCGCTGGCGAGAAACGAGCGGAGCTGCACCAACC 900
DB	841 CCGTGTGAGCGCCAGTGTGGGCGCGCTGGCGAGAAACGAGCGGAGCTGCACCAACC 900
QY	901 CGGCGCTCTCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCCATGACCGCACCGTCT 960
DB	901 CGGCGCTCTCTCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCCATGACCGCACCGTCT 960
QY	961 CCTCTCTGCTTGTCTCTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGGTCCGCTGTG 1020
DB	961 CCTCTCTGCTTGTCTCTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGGTCCGCTGTG 1020
QY	1021 GGTGAGTGTGACCCACTGCGGAGCGCTGAGTGTCTGACCCAGCACCCCGCAACGGAG 1080
DB	1021 GGTGAGTGTGACCCACTGCGGAGCGCTGAGTGTCTGACCCAGCACCCCGCAACGGAG 1080
QY	1081 GGGAGAGTGGCAGGGCACTGACCTGGACACCGCAACTGTACCACTGACCTCTGTGTAC 1140
DB	1081 GGGAGAGTGGCAGGGCACTGACCTGGACACCGCAACTGTACCACTGACCTCTGTGTAC 1140
QY	1141 ACAGTGTCTTGTGCGCTGAGGAGCTGGCCCTCTATGTGGGCTCTATGCGCGTGGCGTCT 1200
DB	1141 ACAGTGTCTTGTGCGCTGAGGAGCTGGCCCTCTATGTGGGCTCTATGCGCGTGGCGTCT 1200
QY	1201 GCTGTGCTGTGCTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
DB	1201 GCTGTGCTGTGCTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
QY	1261 ACTCAGATGTGGTGACTCGTCCATTTCTCAGCTCAGGCTTCCAGCCGCTCAGCATCAAGC 1320
DB	1261 ACTCAGATGTGGTGACTCGTCCATTTCTCAGCTCAGGCTTCCAGCCGCTCAGCATCAAGC 1320
QY	1321 CCAGCAAGCAGACAAACCCCACTCTCTCAACATCCAGCCGAGACCTCAGCACCAACCA 1380
DB	1321 CCAGCAAGCAGACAAACCCCACTCTCTCAACATCCAGCCGAGACCTCAGCACCAACCA 1380
QY	1381 CCTACAGGGCAGTCTGTCTCCCGCAGGATGGGCCAGCCCAAGTTCACGCTCACCA 1440
DB	1381 CCTACAGGGCAGTCTGTCTCCCGCAGGATGGGCCAGCCCAAGTTCACGCTCACCA 1440
QY	1441 ATGGGCACCTGTCTCAGCCCTCGGTGGCGCCGACACACACTGACACCAAGTCTCCCA 1500
DB	1441 ATGGGCACCTGTCTCAGCCCTCGGTGGCGCCGACACACACTGACACCAAGTCTCCCA 1500
QY	1501 CCTCTGAGGCGGAGGAGTGTCTCCGCTCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
DB	1501 CCTCTGAGGCGGAGGAGTGTCTCCGCTCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560
QY	1561 CCGAGGACACCAACATGACCTATGGGACCTTCACTTCTCGGGGGCGGCTGATGA 1620
DB	1561 CCGAGGACACCAACATGACCTATGGGACCTTCACTTCTCGGGGGCGGCTGATGA 1620
QY	1621 TCCTTAATACAGGTATCAGCTCTCTCATCCCGCCAGATGCCATACCCCGAGGGAAGATCT 1680
DB	1621 TCCTTAATACAGGTATCAGCTCTCTCATCCCGCCAGATGCCATACCCCGAGGGAAGATCT 1680
QY	1681 ATCAGATCTACTCAGCTGCAACAGCCGGAAGAGTGGCTTGCCTTACCTAGCTGGCTGC 1740
DB	1681 ATCAGATCTACTCAGCTGCAACAGCCGGAAGAGTGGCTTGCCTTACCTAGCTGGCTGC 1740
QY	1741 AGACCTGTGAGTCCCATGTAGCTGTGACCCCTGGGCTCTCTCTCCTCAGCCGCGAG 1800
DB	1741 AGACCTGTGAGTCCCATGTAGCTGTGACCCCTGGGCTCTCTCTCCTCAGCCGCGAG 1800
QY	1801 TCATCTGTGCTATGGACACTGTGGGAGCCAGCCCTGACAGCTGAGGCTGGCCCTCA 1860
DB	1801 TCATCTGTGCTATGGACACTGTGGGAGCCAGCCCTGACAGCTGAGGCTGGCCCTCA 1860
QY	1861 AAAGAGAGTGTGGAGGGGAGCTGGAGCAGAGTGTGCTGACCTGGGCGAGGAGGCGC 1920
DB	1861 AAAGAGAGTGTGGAGGGGAGCTGGAGCAGAGTGTGCTGACCTGGGCGAGGAGGCGC 1920
QY	1921 CCTCCCACTCTACTTCTGCTGAGCGCAGTGTCTGTCTGTCTTCTTACCCGAGCAGC 1980
DB	1921 CCTCCCACTCTACTTCTGCTGAGCGCAGTGTCTGTCTGTCTTCTTACCCGAGCAGC 1980
QY	1981 TGGGCGCTTTGGCCCTGTGGGAGAGGCGCTTCAAGCTGGGCTGGCGCAAGCGCTCAAGC 2040

Db 1981 TGGGCCCGCTTGGCCCTGGTGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCTCAAGC 2040  
Qy 2041 TGGCTTCTGTTTGGCCCGGTGGCTTGCACCTCCCTCGAGTCAACATCCGGTCTACTGCC 2100  
Db 2041 TGGCTTCTGTTTGGCCCGGTGGCTTGCACCTCCCTCGAGTCAACATCCGGTCTACTGCC 2100  
Qy 2101 TGGATGACACCCAGCATGCACTCAAGGAGTGGTGCAGCTGGAGAGCAGCTGGGGGAC 2160  
Db 2101 TGCATGACACCCAGCATGCACTCAAGGAGTGGTGCAGCTGGAGAGCAGCTGGGGGAC 2160  
Qy 2161 AGCTGATCCAGGAGCCAGCGGTCTGCACTTCAAGGACAGTTACCAAACTTGCCTTAT 2220  
Db 2161 AGCTGATCCAGGAGCCAGCGGTCTGCACTTCAAGGACAGTTACCAAACTTGCCTTAT 2220  
Qy 2221 CCATCCACGATGTGCCCGAGTCCCTGTGGAGAGTAAAGTCCCTGTGAGTACACAGAGA 2280  
Db 2221 CCATCCACGATGTGCCCGAGTCCCTGTGGAGAGTAAAGTCCCTGTGAGTACACAGAGA 2280  
Qy 2281 TCCCTTTTATCATCTGGAATGGACGACGCGGTACTTTGCACTGCACCTTACCCCTGG 2340  
Db 2281 TCCCTTTTATCATCTGGAATGGACGACGCGGTACTTTGCACTGCACCTTACCCCTGG 2340  
Qy 2341 AGCGTGTCAAGGAGTCACTAGTGAACCTGGCTGCAAGCTGTGGTGTGGAGTGGAGG 2400  
Db 2341 AGCGTGTCAAGGAGTCACTAGTGAACCTGGCTGCAAGCTGTGGTGTGGAGTGGAGG 2400  
Qy 2401 GCGACGGGAGAGCTTCAAGCATCAACTTCAACATCAACATCAACATCAACATCAACAT 2460  
Db 2401 GCGACGGGAGAGCTTCAAGCATCAACTTCAACATCAACATCAACATCAACATCAACAT 2460  
Qy 2461 TGTGTGCTCTGGAGAGTGAAGCGGGGTGCCAGCCCTGTGGGCCCCAGTGCCTTCAAGA 2520  
Db 2461 TGTGTGCTCTGGAGAGTGAAGCGGGGTGCCAGCCCTGTGGGCCCCAGTGCCTTCAAGA 2520  
Qy 2521 TCCCTTCTCTCATTCGGCAGAGATATTTCCAGCTGGACCCACCTGTAGCGGGGTG 2580  
Db 2521 TCCCTTCTCTCATTCGGCAGAGATATTTCCAGCTGGACCCACCTGTAGCGGGGTG 2580  
Qy 2581 CCGACTTGGCGGACTCTGGGCCAGAACTCCACCTGGACAGCGCGGCACTTCCGCCAAG 2640  
Db 2581 CCGACTTGGCGGACTCTGGGCCAGAACTCCACCTGGACAGCGCGGCACTTCCGCCAAG 2640  
Qy 2641 CCAAGCCAGCCAGCCAGCCATGATCCTCAACCTGTGGAGAGCGCGGCACTTCCGCCAAG 2700  
Db 2641 CCAAGCCAGCCAGCCAGCCATGATCCTCAACCTGTGGAGAGCGCGGCACTTCCGCCAAG 2700  
Qy 2701 GCAACCTCAGCAGCTGGCTGAGCAGTGGTGGACTGGGCGAGCCAGCGTGGCTCT 2760  
Db 2701 GCAACCTCAGCAGCTGGCTGAGCAGTGGTGGACTGGGCGAGCCAGCGTGGCTCT 2760  
Qy 2761 TCACAGTGTCCGAGGTGAGTGTGAGGCGCGGCGGCGGCGGCACTTCACTTCAACAG 2820  
Db 2761 TCACAGTGTCCGAGGTGAGTGTGAGGCGCGGCGGCGGCGGCACTTCACTTCAACAG 2820  
Qy 2821 CTTTGGACCCCAAGGACAGCGAGAACCGGACAGGGGCGCTTCCGCCACACCGGGAG 2880  
Db 2821 CTTTGGACCCCAAGGACAGCGAGAACCGGACAGGGGCGCTTCCGCCACACCGGGAG 2880  
Qy 2881 A 2881  
Db 2881 A 2881

RESULT 2  
LOCUS AX449572  
DEFINITION Sequence 1 from Patent WO0210216.  
ACCESSION AX449572  
VERSION AX449572.1 GI:21698195  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Padigaru, M., Mezes, P., Mishra, V., Burgess, C., Casman, S.,  
Grosse, W.M., Alsebrook, J.P., Lepley, D.M., Gerlach, V.L.,  
Macdougall, J.R. and Smithson, G.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 0210216-A 1 07-FEB-2002;  
Curagen Corporation (US)  
FEATURES  
source 1. 2752  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 92.9%; Score 2676.4; DB 6; Length 2752;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 2728; Conservative 0; Mismatches 21; Indels 9; Gaps 3;  
Qy 42 CCGCGGGGCCCCCGCGCCCGCCCGCTGCGCCGCGCCGCGCCATGGCCGCTCCGGCCC 101  
Db 1 CCGCGGGGCCCCCGCGCCCGCCCGCTGCGCCGCGCCGCGCCATGGCCGCTCCGGCCC 60  
Qy 102 GGCCTGTGGCCAGCGCTCTGGGCATAGTCTCGCCGCTTGGCTTCGCGGCTCGGGTGC 161  
Db 61 GGCCTGTGGCCAGCGCTCTGGGCATAGTCTCGCCGCTTGGCTTCGCGGCTCGGGTGC 120  
Qy 162 CAGCAGAGTGGCCACCCTGGCCAAACCCAGTGCCTGGTGCACAAACCCGACCTGCTCCCCAC 221  
Db 121 CAGCAGAGTGGCCACCCTGGCCAAACCCAGTGCCTGGTGCACAAACCCGACCTGCTCCCCAC 180  
Qy 222 TTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAAGCCAGTGTCTTGTGTGC 281  
Db 181 TTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAAGCCAGTGTCTTGTGTGC 240  
Qy 282 AAGGCGCTGCCCGCCAGCAGATCTTCTCAAGTGAACGGGAGTGGGTGGCCAGGTG 341  
Db 241 AAGGCGCTGCCCGCCAGCAGATCTTCTCAAGTGAACGGGAGTGGGTGGCCAGGTG 300  
Qy 342 GACCACGTGATCGAGCGCAGCAGACGAGCGAGTGTGAGCCGACCATGGAGGTCCGC 401  
Db 301 GACCACGTGATCGAGCGCAGCAGACGAGCGAGTGTGAGCCGACCATGGAGGTCCGC 360  
Qy 402 ATTAATGTCTAAGCGCAGCAGTTCGAGAAAGTGTTCGGGCTGGAGGAATACCTGTCCTAG 461  
Db 361 ATTAATGTCTAAGCGCAGCAGTTCGAGAAAGTGTTCGGGCTGGAGGAATACCTGTCCTAG 420  
Qy 462 TCGGTGGCATGGAGTCTTCGGGCGACCAAGAGTCAAGAGTCAAGAGCTTACATCCGCATAGCC 521  
Db 421 TCGGTGGCATGGAGTCTTCGGGCGACCAAGAGTCAAGAGTCAAGAGCTTACATCCGCATAGCC 480  
Qy 522 AGATTGGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAGGAGGTGTCCCTGGAGCAGGGC 581  
Db 481 AGATTGGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAGGAGGTGTCCCTGGAGCAGGGC 540  
Qy 582 ATCTGCTGCCCTGCCCTCCACCGGAGGCGCATCCCTCCAGCCGAGTGGAGTGGGTCCGG 641  
Db 541 ATCTGCTGCCCTGCCCTCCACCGGAGGCGCATCCCTCCAGCCGAGTGGAGTGGGTCCGG 600  
Qy 642 AACGAGGACCTGGTGGACCGCTTGTGACCGCCCAATGTATACATCACGCGGAGCACAGC 701  
Db 601 AACGAGGACCTGGTGGACCGCTTGTGACCGCCCAATGTATACATCACGCGGAGCACAGC 660  
Qy 702 CTGTGGTGGCAGCGCCCGCTTGTGACCGCCCAATGTATACATCACGCGGAGCACAGC 761  
Db 661 CTGTGGTGGCAGCGCCCGCTTGTGACCGCCCAATGTATACATCACGCGGAGCACAGC 720  
Qy 762 ATCTGGCAGCTCGCGCGCAGCGCTTCCGCTGTCTACATCGTCAAGAGTGGGTGG 821  
Db 721 ATCTGGCAGCTCGCGCGCAGCGCTTCCGCTGTCTACATCGTCAAGAGTGGGTGG 780  
Qy 822 TCGAGTGGACCGAGTGGTCCGTGTGAGCGCCAGCTGTGGGCGCGCTGGCAGAAACGG 881

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

Db	781	TCGACGTGACCGAGTGGTCCGCTCTGACGGCCAGCTGTGGCGCGGTGGCAGAAA	CGG	840
QY	882	AGCGGAGCTGCACAAACCCGGGCGCTCTCAACGGGGGGCGCTTTCTGTGAGGGGCAGAA	T	941
Db	841	AGCGGAGCTGCACCAACCCGGCGCTCTCAACGGGGGGCGCTTTCTGTGAGGGGCAGAA	T	900
QY	942	GTCCATGACCGACCGCTCTCTCTGCTGTGTCGTGCGACGGCAGCTGGAGCCCGTGG	1001	
Db	901	GTCCAGAA---AAACAGCCTGGCCACCCCTGTGCCAGTACGCGAGCTGGAGCCCGTGG	957	
QY	1002	AGCAAGTGGTCGGGCTGTGGGCTTGGACTGCACCCACTGGCGAGCGGTGAGTCTCTGAC	1061	
Db	958	AGCAAGTGGTCGGGCTGTGGGCTTGGACTGCACCCACTGGCGAGCGGTGAGTCTCTGAC	1017	
QY	1062	CCAGACCCCGCGAAACGAGGGGAGGAGTGCAGGGCACTGACCTGGACACCCGCAACTGT	1121	
Db	1018	CCAGACCCCGCGCAACGAGGGGAGGAGTGCAGGGCACTGACCTGGACACCCGCAACTGT	1077	
QY	1122	ACCAGTGACCTCTGTGTACACAGTCTTCTGGGCCCTGAGACGTGGCCCTCTATGTGGGC	1181	
Db	1078	ACCAGTGACCTCTGTGTACACAGTCTTCTGGGCCCTGAGACGTGGCCCTCTATGTGGGC	1137	
QY	1182	CTCATCCGCGTGGCGGCTCTGCTGGTCCCTGCTCTGCTTGTCTCTCATCTCGCTTATTGC	1241	
Db	1138	CTCATCCGCGTGGCGGCTCTGCTGGTCCCTGCTCTGCTCTGCTCTCATCTCGCTTATTGC	1197	
QY	1242	CGGAAGAGGAGGGGCTGGACTCAGATGTGGCTGACTGTGTCATTTCTCACTCAGGCTTC	1301	
Db	1198	CGGAAGAGGAGGGGCTGGACTCAGATGTGGCTGACTGTGTCATTTCTCACTCAGGCTTC	1257	
QY	1302	CAGCCGCTCAGCATCAAGCCAGCAAGCAGACAAACCCCATCTGCTCAACATCCAGCCG	1361	
Db	1258	CAGCCGCTCAGCATCAAGCCAGCAAGCAGACAAACCCCATCTGCTCAACATCCAGCCG	1317	
QY	1362	GACCTTCAG---CACCACACCACTACCAAGGCGAGTCTGTCTCCCGCAGGATGGCC	1418	
Db	1318	GACCTCAGCACCAACCAACCACTACCAAGGCGAGTCTGTCTCCCGCAGGATGGCC	1377	
QY	1419	AGCCCCAAGTTCAGCTCAACAATGGGGA	1478	
Db	1378	AGCCCCAAGTTCAGCTCAACAATGGGGA	1437	
QY	1479	ACACTGCACACAGCTCTCCACCTCTGAGSCGAGGAGTTCGTCTCCCGCCTCTCCACC	1538	
Db	1438	ACACTGCACACAGCTCTCCACCTCTGAGSCGAGGAGTTCGTCTCCCGCCTCTCCACC	1497	
QY	1539	CAGAACTACTTCCGCTCCCTGCCCGGACCAAGCAACATGACCTATGGGACCTTCAAC	1598	
Db	1498	CAGAACTACTTCCGCTCCCTGCCCGGACCAAGCAACATGACCTATGGGACCTTCAAC	1557	
QY	1599	TTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCATCTCCCCCAGAT	1658	
Db	1558	TTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCATCTCCCCCAGAT	1617	
QY	1659	GCCATACCCGAGGAAAGATCTATGAGATCTACCTCAGCTGCAACAAGCCGGAAGACGTG	1718	
Db	1618	GCCATACCCGAGGAAAGATCTATGAGATCTACCTCAGCTGCAACAAGCCGGAAGACGTG	1677	
QY	1719	AGGTTGCCCTAGCTGGGCTGTCCAGACCCCTGCTGAGTCCCATCGTTAGCTGTGAGACCCCT	1778	
Db	1678	AGGTTGCCCTAGCTGGGCTGTCCAGACCCCTGCTGAGTCCCATCGTTAGCTGTGAGACCCCT	1737	
QY	1779	GGGCTCTGCTCACCGGCGAGTCACTCTGGCTATGGACCACTCTGGGGAGCCAGACCT	1838	
Db	1738	GGGCTCTGCTCACCGGCGAGTCACTCTGGCTATGGACCACTCTGGGGAGCCAGACCT	1797	
QY	1839	GACAGTGAGCCTGCGCTCAAAAAGCAGTCTGTCGAGGCGAGCTGGGAGCAGGATGTG	1898	
Db	1798	GACAGTGAGCCTGCGCTCAAAAAGCAGTCTGTCGAGGCGAGCTGGGAGCAGGATGTG	1854	
QY	1899	CTGCACCTGGGAGGAGGCGCCCTCCACCTCTACTGCGCAGCTGGAGCCAGTGCC	1958	

Db	1855	CTGCACCTGGCGAGAGGGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC	191
Qy	1959	TGCTACGTTCTTACCAGCAGAGCTGGCGCGCTTTGGCCCTGGTGGGAGAGGCCCTCAGCGTG	2018
Db	1915	TGCTACGTTCTTACCAGCAGAGCTGGCGCGCTTTGGCCCTGGTGGGAGAGGCCCTCAGCGTG	1974
Qy	2019	GCTGGCCGCAAGCGCCTCAAGCTGCTTCTGTTTGGCCGCGGTGGCTGCACCTCCCTCGAG	2078
Db	1975	GCTGGCCGCAAGCGCCTCAAGCTGCTTCTGTTTGGCCGCGGTGGCTGCACCTCCCTCGAG	2034
Qy	2079	TACAAATCATCGGCTTACTGTCCTGTCATGACACCCACGATGCACTCAAGGAGGTTGGTGCAG	2138
Db	2035	TACAAATCATCGGCTTACTGTCCTGTCATGACACCCACGATGCACTCAAGGAGGTTGGTGCAG	2094
Qy	2139	CTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCACAAGGTCCTGCACTTCAAGGAC	2198
Db	2095	CTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCACAAGGTCCTGCACTTCAAGGAC	2154
Qy	2199	AGTTTACCAAACTTGGCCCTATTCATCCACGATGTGCCAGCTCCCTGTGGAGAGTAAG	2258
Db	2155	AGTTTACCAAACTTGGCCCTATTCATCCACGATGTGCCAGCTCCCTGTGGAGAGTAAG	2214
Qy	2259	CTCCTTGTACGTACCAAGGAGATCCCTCTTTTATCACATCTGGAATGCACGACGCGGTAC	2318
Db	2215	CTCCTTGTACGTACCAAGGAGATCCCTCTTTTATCACATCTGGAATGCACGACGCGGTAC	2274
Qy	2319	TTGCACCTGCACCTTTCACCCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCTGTGAAG	2378
Db	2275	TTGCACCTGCACCTTTCACCCCTGGAGCGTGTGAGCCCCAGCACTAGTGACCTGGCTGTGAAG	2334
Qy	2379	CTGTGGGTGTGGCAGGTGTGGAGGGGACGGGAGAGCTTCAGATCAACTTCAACATCACC	2438
Db	2335	CTGTGGGTGTGGCAGGTGTGGAGGGGACGGGAGAGCTTCAGATCAACTTCAACATCACC	2394
Qy	2439	AAGACACAAAGTTTGTCTGAGCTCTGGCTCTGGAGAGTGAAGCGGGGTCCCAAGCCCTG	2498
Db	2395	AAGACACAAAGTTTGTCTGAGCTCTGGCTCTGGAGAGTGAAGCGGGGTCCCAAGCCCTG	2454
Qy	2499	GTGGGCCCCAGTGCTTCAAGATCCCTTCTCTATTCCGCAAGAAGATAATTTCCAGCCTG	2558
Db	2455	GTGGGCCCCAGTGCTTCAAGATCCCTTCTCTATTCCGCAAGAAGATAATTTCCAGCCTG	2514
Qy	2559	GACCAACCTTAGCGGGGTGCGACTGGCGGACTCTGGCCGAGAACTCCACCTGGAC	2618
Db	2515	GACCAACCTTAGCGGGGTGCGACTGGCGGACTCTGGCCGAGAACTCCACCTGGAC	2574
Qy	2619	AGCCATCTCAGCTTCTTTGGCTCCAAAGCCAGCCCAAGCCATGATCCTCAACCTGTGG	2678
Db	2575	AGCCATCTCAGCTTCTTTGGCTCCAAAGCCAGCCCAAGCCATGATCCTCAACCTGTGG	2634
Qy	2679	GAGGCGCGCACTTCCCAACCGCAACTCAGCCAGCTGGCTGACAGAGTGGCTGGACTG	2738
Db	2635	GAGGCGCGCACTTCCCAACCGCAACTCAGCCAGCTGGCTGACAGAGTGGCTGGACTG	2699
Qy	2739	GCCAGCCAGACGCTGGCTCTTACAGTGTGAGGGCTGAGTGTGAGGCGGCGCAG	2796
Db	2695	GCCAGCCAGACGCTGGCTCTTACAGTGTGAGGGCTGAGTGTGAGGCGGCGCAG	2752
RESULT 3			
AX451652			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1.			
Koehler, R.H.			
AUTHORS			
REGULATION OF human netrin binding membrane receptor unc5h-1			
TITLE			
AX451652			
Sequence 1 from Patent WO0233080.			
AX451652.1			
GI:21698587			
2697 bp			
DNA			
linear			
PAT 03-JUL-2000			



[illegible][illegible]

TITLE	Receptors
JOURNAL	Patent: WO 0198354-A 13 27-DEC-2001;
FEATURES	Incyte Genomics, Inc. (US) Location/Qualifiers

```

source      1. .3580
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="Incyte ID No: 6052371CB1"

ORIGIN
Query Match      83.4%; Score 2402.6; DB 6; Length 3580;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;

QY      84  GCCATGACCGCTGCGGCGCCGGCCCTGTGGCCACAGCGCTCTCTGGGCATAGTCTCTCGCGCGTTGG 143
Db       1  GCCATGACCGCTGCGGCGCCGGCCCTGTGGCCACAGCGCTCTCTGGGCATAGTCTCTCGCGCGTTGG 60

QY     144  CTCGCGCGCTCGGGTGCACAGCAGTAGTGCCACCGTGGCCCAACCCAGTGCCCTGTGTGGCCCAAC 203
Db      61  CTCGCGCGCTCGGGTGCACAGCAGTAGTGCCACCGTGGCCCAACCCAGTGCCCTGTGTGGCCCAAC 120

QY     204  CCGGACCTGCTTCCCACTTCTCTGTGGAGCCCGGAGGATGTGTACATCGTCTCAAGAAACAAG 263
Db     121  CCGGACCTGCTTCCCACTTCTCTGTGGAGCCCGGAGGATGTGTACATCGTCTCAAGAAACAAG 180

QY     264  CCAGTGTCTGTGTGTGTCACAGGCGGTGCCCGCACGCGAGATCTTCTTCAAGTGGAAACGG 323
Db     181  CCAGTGTCTGTGTGTGTCACAGGCGGTGCCCGCACGCGAGATCTTCTTCAAGTGGCAACGG 240

QY     324  GAGTGGGTGCGCAGGTGGACCAAGTATCGAGCGCAGCACAGACGGGAGCAGTGTGTGAG 383
Db     241  GAGTGGGTGCGCAGGTGGACCAAGTATCGAGCGCAGCACAGACGGGAGCAGTGTGTGAG 300

QY     384  CCGACCATGAGGTCGCGCATTAATGTCTCAAGGCAGCAGGTTCGAGAAAGTGTTCGGGCTG 443
Db     301  CCCACCATGAGGTCGCGCATTAATGTCTCAAGGCAGCAGGTTCGAGAAAGTGTTCGGGCTG 360

QY     444  GAGGAATACTGGTGCCAGTGCGGTGGCATGAGACTCTCGGGCACCACCAAGACTCACAAG 503
Db     361  GAGGAATACTGGTGCCAGTGCGGTGGCATGAGACTCTCGGGCACCACCAAGAGTTCAGAAG 420

QY     504  GCCTACATCCGCATAGCCAGATTTCGCAAGAACTTCGAGCAGGACGCGCTGGCCCAAGAG 563
Db     421  GCCTACATCCGCATAGCCATTTGGCAAGAACTTCGAGCAGGACGCGCTGGCCCAAGAG 480

QY     564  GTGTCTCTGTGAGCAGGGCATCGTCTGCCTGCGTCCACGAGGGGATCCCTTCAGACC 623
Db     481  GTGTCTCTGTGAGCAGGGCATCGTCTGCCTGCGTCCACGAGGGGATCCCTTCAGACC 540

QY     624  GAGGTGGAGTGGCTCCGAAACGAGGACTGTGTGACCCGCTCCCTGGACCCCAAGTGTATAC 683
Db     541  GAGGTGGAGTGGCTCCGAAACGAGGACTGTGTGACCCGCTCCCTGGACCCCAAGTGTATAC 600

QY     684  ATCAGCGGGAGCAGAGCTGTGTGCGACAGGCGCGCTTGTGTGACAGCGGCAACTAC 743
Db     601  ATCAGCGGGAGCAGAGCTGTGTGCGACAGGCGCGCTTGTGTGACAGCGGCAACTAC 660

QY     744  ACCTCGTGGCCAGAAACATCGTGGCAGCTGCCGACGCGCTCGGTGTGTGTATCATGTC 803
Db     661  ACCTCGTGGCCAGAAACATCGTGGCAGCTGCCGACGCGCTCGGTGTGTGTATCATGTC 720

QY     804  TAGGTGAACGGTGGGTGGTGCAGCTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGG 863
Db     721  TAC----- 723

QY     864  CCGCGCTGCGAAGAACGAGCGCGGAGTGCACCAACCGCGCGCTCTCAACGCGGGCGCT 923
Db     724  ----- 723

QY     924  TTCTGTGAGGGGCAGAAATGTCCATGACCGACCGTCTCTCTCTGTCTTGTCTCTGTGAC 983
Db     724  -----GTGGAC 729

QY     984  GGCAGCTGGAGCCCGTGGAGCAAGTGGTGGCGCTGGGCTGGAGTGGAGCTGCAACCACTGGCGG 1043

```



Db 730 GGAGCTGGAGCCCGTGGAGCAAGTGTGGCGCTGTGGCTGGACTGACACCACTGGCGG 789  
Qy 1044 AGCCGTGAGTGTCTGTGACCCAGCACCCTGCAAGCGAGGGAGGAGTGCAGGGCACTGAC 1103  
Db 790 AGCCGTGAGTGTCTGTGACCCAGCACCCTGCAAGCGAGGGAGGAGTGCAGGGCACTGAC 849  
Qy 1104 CTGAGCACCCGCAACTGTGACAGTGAACCTCTGTGTACACAGTGTCTGTGGCCCTGAGGAC 1163  
Db 850 CTGAGCACCCGCAACTGTGACAGTGAACCTCTGTGTACACAGTGTCTGTGGCCCTGAGGAC 909  
Qy 1164 GTGGCCCTCTATGTGTGGCCCTCATGCGCGTGGCGCTGTGCTGGTCTGTGCTGTGTGTC 1223  
Db 910 GTGGCCCTCTATGTGTGGCCCTCATGCGCGTGGCGCTGTGCTGGTCTGTGCTGTGTGTC 969  
Qy 1224 CTATCCTCTGTTATTGTCGGAAGAAGAGAGGGGTGGACTCAGATGTGGCTGACTCGTCC 1283  
Db 970 CTATCCTCTGTTATTGTCGGAAGAAGAGAGGGGTGGACTCAGATGTGGCTGACTCGTCC 1029  
Qy 1284 ATTCTACCTCAGGCTTCCAGCCCGTCCAGCATCAAGGCCAGCAAGAGCAGACACCCCAT 1343  
Db 1030 ATTCTACCTCAGGCTTCCAGCCCGTCCAGCATCAAGGCCAGCAAGAGCAGACACCCCAT 1089  
Qy 1344 CTGCTACCATCCAGCGGAGCTCAG---CACACACACAGCTTACAGAGGAGTCTGTGT 1400  
Db 1090 CTGCTACCATCCAGCGGAGCTCAGCACCACCAACCACTTACAGGGAGTCTGTGT 1149  
Qy 1401 CCCGGCAGGATGGGCCCGCAGCCCAAGTTCAGCTCACCAGTGGCAGCTGTGCTCAGCCCC 1460  
Db 1150 CCCGGCAGGATGGGCCCGCAGCCCAAGTTCAGCTCACCAGTGGCAGCTGTGCTCAGCCCC 1209  
Qy 1461 CTGGTGGCGGCGGCGCACACACTGCACACAGTCTCCACCTCTGAGGCGGAGGAGTTC 1520  
Db 1210 CTGGTGGCGGCGGCGCACACACTGCACACAGTCTCCACCTCTGAGGCGGAGGAGTTC 1269  
Qy 1521 GTCTCCGCTCTCCACCCAGACTTCTCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1580  
Db 1270 GTCTCCGCTCTCCACCCAGACTTCTCGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1329  
Qy 1581 ACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTATACAGTATCAGC 1640  
Db 1330 ACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTATACAGTATCAGC 1389  
Qy 1641 CTCTCATCCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACCTCAGCTG 1700  
Db 1390 CTCTCATCCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACCTCAGCTG 1449  
Qy 1701 CACAAGCGGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760  
Db 1450 CACAAGCGGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509  
Qy 1761 GTTAGTGTGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1820  
Db 1510 GTTAGTGTGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1569  
Qy 1821 TGTGGGAGCCAGCCCTGACAGCTGAGCTGCGCTCAAAAGCAGTCTGTCGAGGGG 1880  
Db 1570 TGTGGGAGCCAGCCCTGACAGCTGAGCTGCGCTCAAAAGCAGTCTGTCGAGGGG 1629  
Qy 1881 AGCTGGAGCAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940  
Db 1630 AGCTGGG---AGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686  
Qy 1941 CAGCTGGAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2000  
Db 1687 CAGCTGGAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746  
Qy 2001 GGAGAGCCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2060  
Db 1747 GGAGAGCCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1806  
Qy 2061 GCTGTGACCTCCTCAGTGTACATCCCGGCTTACTGCTGATGACACCAAGATGCA 2120  
Db 1807 GCTGTGACCTCCTCAGTGTACATCCCGGCTTACTGCTGATGACACCAAGATGCA 1866

Qy 2121 CTCAAGAGAGTGTGTGAGTGGAGAGCAGCTGGGGGACAGCTGATCCAGAGCCACGG 2180  
Db 1867 CTCAAGAGAGTGTGTGAGTGGAGAGCAGCTGGGGGACAGCTGATCCAGAGCCACGG 1926  
Qy 2181 GTCTGTGACATTCAGGACAGTTCACCAACCTGCGCTTATCCATCCAGAGTGTGCCACAGC 2240  
Db 1927 GTCTGTGACATTCAGGACAGTTCACCAACCTGCGCTTATCCATCCAGAGTGTGCCACAGC 1986  
Qy 2241 TCCCTGTGGAGAGTAAGTCTCTTGTGAGTCAACAGGAGATCCCTTTTATCATCTGG 2300  
Db 1987 TCCCTGTGGAGAGTAAGTCTCTTGTGAGTCAACAGGAGATCCCTTTTATCATCTGG 2046  
Qy 2301 AATGGCAGCAGCGGTACTTGCACTGCACTTCAACCTTGGAGCGTGTGAGCGGAGCTTCAAGC 2360  
Db 2047 AATGGCAGCAGCGGTACTTGCACTGCACTTCAACCTTGGAGCGTGTGAGCGGAGCTTCAAGC 2106  
Qy 2361 AGTGACCTGCGCTGCAAGCTGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2420  
Db 2107 AGTGACCTGCGCTGCAAGCTGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2166  
Qy 2421 ATCAACTTCAACATCACCAAGGACACAAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGAGTGTGAG 2480  
Db 2167 ATCAACTTCAACATCACCAAGGACACAAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGAGTGTGAG 2226  
Qy 2481 GCGGGGTCCAGCCCTGCTGGTGGGCGGCGGCTTCAAGATCCCTTCTCATTCGGCAG 2540  
Db 2227 GCGGGGTCCAGCCCTGCTGGTGGGCGGCGGCTTCAAGATCCCTTCTCATTCGGCAG 2286  
Qy 2541 AAGATTAATTTCCAGCTTGGACCCACCTGAGGGGGTGGCGGCTGCGGAGTGTGCGG 2600  
Db 2287 AAGATTAATTTCCAGCTTGGACCCACCTGAGGGGGTGGCGGCTGCGGAGTGTGCGG 2346  
Qy 2601 CAGAAATCTCAGCTTGGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCGCCACAGCC 2660  
Db 2347 CAGAAATCTCAGCTTGGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCGCCACAGCC 2406  
Qy 2661 ATGATCTCAACTGTGGAGGCGGCGCTTCCCAAGCCAGCCGCCACAGCCGCCACAGCC 2720  
Db 2407 ATGATCTCAACTGTGGAGGCGGCGCTTCCCAAGCCAGCCGCCACAGCCGCCACAGCC 2466  
Qy 2721 GCAGCAGTGTGCTGAGTGTGGCGCAGCAGCTGCGGCTTCTTACAGTGTGCGAGGCTGAG 2780  
Db 2467 GCAGCAGTGTGCTGAGTGTGGCGCAGCAGCTGCGGCTTCTTACAGTGTGCGAGGCTGAG 2526  
Qy 2781 TGTGTAGGCGGCGCAGCAGCAGCTTACACTCTCAGCAGCTTGTGACCCACCAAGGAC 2840  
Db 2527 TGTGTAGGCGGCGCAGCAGCAGCTTACACTCTCAGCAGCTTGTGACCCACCAAGGAC 2586  
Qy 2841 AGGCAGAGCCGACAGGCGGCTTCCCAAGCCAGCCGCCAGGAG 2881  
Db 2587 AGGCAGAGCCGACAGGCGGCTTCCCAAGCCAGCCGCCAGGAG 2627

## RESULT 5

MMU487852

LOCUS

DEFINITION

AJ487852

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

1

Engelkamp,D.

Cloning of three mouse Unc5 genes and their expression patterns at

mid-gestation

Mech. Dev. 118 (1-2), 191-197 (2002)

12351186

2 (bases 1 to 3992)

MMU487852 3992 bp mRNA linear ROD 24-SEP-2002  
Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).

AJ487852.1 GI:22035783  
netrin receptor Unc5h1; Unc5h1 gene.  
Mus musculus (house mouse)

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1  
Engelkamp,D.  
Cloning of three mouse Unc5 genes and their expression patterns at  
mid-gestation  
Mech. Dev. 118 (1-2), 191-197 (2002)  
12351186  
2 (bases 1 to 3992)

AUTHORS Engelkamp, D.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for Brain Research, Deutscherordenstrasse 46, Frankfurt 60528, GERMANY

FEATURES  
 source Location/Qualifiers  
 1..3992  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 1..3992  
 /gene="Unc5h1"  
 232..2928  
 /gene="Unc5h1"  
 /codon\_start=1  
 /product="netrin receptor Unc5h1"  
 /protein\_id="CAD32250.1"  
 /db\_xref="GI:22035784"  
 /db\_xref="GOA:O8K154"  
 /db\_xref="SPTREMBL:O8K154"  
 /translation="MAVRPLGLWALLIGVLTWLRGSAQOQSATVANPVPANPDLLP  
 HFLVPEDEVIVKKNPVLLVCAVPATQIFKCNGEWVRQDVHIERSDIGSGSLPTM  
 EVRINYSRQVQVFLGLEYWCQVAMSSGVTQSKQAYIRIAYLRKNFQEPLEKEV  
 SLRQGVLPQRPPEGIPPAEVEWLRNEDLDVDPNDNVIYTRHSLVVRQARLADTAN  
 YTCVAKNIVARRESAAAVIVVNGWSWTWESVCSACGCRQKRSRSTNPAPLN  
 GGAPCGQNVQKATCLCPVDGSMSPMSKAGLDCDTHWRSEKSDPAPRNGBEC  
 RGADLDRNCTSLCLHTSSGPEDVALYIGLVAVAVCLILLVLLVLYCRKEGLDS  
 DVADSLTSGFQFVSIKPSKADNPILLITIQDLSITTTTQSLCPRQGPSPKQL  
 SNHLLSPLGSGRHTLHSSPTSEAEFVSRSTONYFRSLPRGTSMAYGTENFELGG  
 RLMPNTGISLLIPPAIPRKIVIEYILTKHPEDVRLPLAGCQTLSPVSCGPGV  
 LLTRPVITLMDHGEPRDPSWLRKQSCGSEWEDVHLGESSPHLYYCOLEAGAC  
 YVTEQLGRPALVGEALSVATKRLRLLPAPVACTSLVKNIRVYCLHDTHDALKEV  
 QLEKQLGQLIQRVRLFKDSYHNLRSLSDHPVPSLWKSLLVSVQIEIPFYHWNKT  
 QOYLHCTFTLRVNVASLDAKQVWQVDEGDSFNINFTNDRFAEMLALESBG  
 GVPLVGPFSAPKIPFLIRQKIITSLDPPCSRGDWRTLAQKLHDLHSLSPFASKPSPT  
 AMTLNLWEARHFPNGLQGLAAAVALGLQPDAGLFTVSEAE"

ORIGIN

Query Match 80.8%; Score 2327.8; DB 10; Length 3992;  
 Best Local Similarity 89.3%; Pred. No. 0;  
 Matches 2579; Conservative 0; Mismatches 292; Indels 16; Gaps 6;

QY 1 AGCTGGGCTCCGGCTGAGCGGT-AAAGCGCGCTCCCGCGCGGGGCGCGCGCC 59  
 DB 149 AGGTGGGCTCTCGGTGAGCGGTGACAGCTCTCTCCCGCGCGGGGCGCTAGTCC 208  
 QY 60 GGGCCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 119  
 DB 209 AGCCCGCTCGT---CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 264  
 QY 120 CTGGGCTAGTCTCGCGCGCTGCTCGCGCGCTCGCGCGCGCGCGCGCGCGCGCTG 179  
 DB 265 CTGGGCTAGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 324  
 QY 180 GCCAACCGAGTGGTGGCAACCGCGCGCTGCTCCCGCGCGCGCGCGCGCGCGCG 239  
 DB 325 GCCAACCGAGTGGTGGCAACCGCGCGCTGCTCCCGCGCGCGCGCGCGCGCGCG 384  
 QY 240 GATGTGATCGTCAAGAAACAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 299  
 DB 385 GACGTGTACATTGTCAAGAAACAGCGCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGT 444  
 QY 300 CAGATCTTCTCAAGTGCAACGGGAGTGGGTGGCGCGCGCGCGCGCGCGCGCGCGCG 359  
 DB 445 CAGATCTTCTCAAGTGCAACGGGAGTGGGTGGCGCGCGCGCGCGCGCGCGCGCGCG 504  
 QY 360 AGCAGACGGGAGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419  
 DB 505 AGCAGTGGGAGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 564  
 QY 420 CAGTGTGAGAGGTGTTGGGCTGGAGGAATACTGGTGGCGAGTGGTGGCGCGCGCGCG 479  
 DB 565 CAGTGTGAGAGGTGTTGGGCTGGAGGAGTACTGGTGGCGAGTGGTGGCGCGCGCGCG 624

QY 480 TGGGGCACCAAGAGTCAAGAGGCTACATCCGATAGCAGATTGGCAAGAACTTC 539  
 DB 625 TCAGGAAACACCAAAAGCCAGAGGCTACATCCGATTGCCTATTTCGCAAGAACTTT 684  
 QY 540 GAGCAGGAGCCGTGGCCAAAGAGGTGTCTCTGGAGCAGGCGATCGTGTGCGCTGCGGT 599  
 DB 685 GAGCAGGAGCCGTGGCCAAAGAGGTGTCTCTGGAGCAGGCGATTTGTCTACCTTGTGCG 744  
 QY 600 CCACCGAGGGGATCCCTCCAGCCGAGGTGGGTCTCGGAACAGAGACCTTGTGTGAC 659  
 DB 745 CCCCCGAGGAATCCCCCAGCTGAGGTGGGTGGCTCGAAATAGAGACCTCTGTGGAC 804  
 QY 660 CCGTCTGTGACCGCCCAATGTATACATCACGCGGAGCACAGCCTGTGTGTGCGACAGCC 719  
 DB 805 CCGTCTGTGACCGCCCAATGTGTATACATCACGCGGAGCACAGCCTGTGTGTGCGAGCC 864  
 QY 720 CGCTTGTGTGACCGCCCAATGTATACATCACGCGGAGCACAGCCTGTGTGTGCGACAGCC 779  
 DB 865 CGCTTGTGTGACCGCCCAATGTATACATCACGCGGAGCACAGCCTGTGTGTGCGAGCC 924  
 QY 780 AGCGCTCTCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 839  
 DB 925 AGCGCTCTCGCGCGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 984  
 QY 840 TCGCTGTGACCGCCCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899  
 DB 985 TCGCTGTGACCGCCCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1044  
 QY 900 CCGCGCTCTCAACGGGCGCGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 959  
 DB 1045 CCGCGCTCTCAACGGGCGCGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1104  
 QY 960 TCCTCTGTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1019  
 DB 1105 GCCACTCTG---TGCCAGTGGATGGAGCTGGAGCCATGGAGTAAGTGTGTGTGTGTGTGT 1161  
 QY 1020 GGGCTGAGTGCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGGA 1079  
 DB 1162 GGGCTTGTGCTGACCCCACTGGCGGAGCGCGGAGTGTCTCCAGCCAGCGCGCGCAACGGA 1221  
 QY 1080 GGGGAGAGTGCACAGGCGACTGACCTGGACACCGCGAACTGTACAGTGTGTGTGTGTGTGT 1139  
 DB 1222 GGTGAGGAGTGCAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1281  
 QY 1140 CACAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1199  
 DB 1282 CACACTCTTTCGCGCGCGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1341  
 QY 1200 TGGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1259  
 DB 1342 TGGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1401  
 QY 1260 GACTCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1319  
 DB 1402 GACTCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1461  
 QY 1320 CCCAGCAAGCAGACAAACCCCATCTGTCTACCATCCAGCCGAGCCTCAGCACCAAC--- 1376  
 DB 1462 CCCAGCAAGCAGACAAATCCCATCTGTCTACCATCCAAACCGGAGCTCAGCACCAACGAG 1521  
 QY 1377 ACCACTTACAGGCGAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1436  
 DB 1522 ACCACTTACAGGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1581  
 QY 1437 ACCATGGGACCTGTCTCAGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1496  
 DB 1582 TCTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1641  
 QY 1497 CCCACCTCTGAGGCGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1556  
 DB 1642 CCCACCTCTGAGGCTGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1701  
 QY 1557 CTGCCCCGAGGACACAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGGCGCGGTGT 1616

```
Db 1702 CTGCCCCGGGTACAGCAACATGGCCTATGGGACCTTCAACTTCTCGGGGGCGGGTG 1761
QY 1617 ATGATCCCTAATACAGGTATCAGCTCCTCATCCCCCAGATGCCATACCCCGAGGAAG 1676
Db 1762 ATGATCCCTAATACAGGAATCAGCTCCTCATATCCCCCGGACGCAATCCCCCGAGGAAG 1821
QY 1677 ATCTATGAGATCTACCTCAGCTGCACAAGCGGGAAGACGTGAGGTTGGCCCTAGCTGGC 1736
Db 1822 ATCTACGAGATCTACCTCACTCTGCACAGCCAGAGACGTGAGGTTGGCCCTAGCTGGC 1881
QY 1737 TGTACAGCCCTGTGAGTCCCATCGTTAGTGTGGAACCCCTGTGCGCTCTGCTCACCCGG 1796
Db 1882 TGTACAGCCCTGTGAGTCCCATCGTTAGTGTGGAACCCCTGTGCGCTCTGCTCACCCGG 1941
QY 1797 CCAGTCACTCGGTATGAGCACTGTGGGAGGCCAGCCCTGACAGCTGAGGCTCGGC 1856
Db 1942 CCAGTCACTCGGTATGAGCACTGTGGGAGGCCAGCCCTGAGGAGCCAGCTCGGC 2001
QY 1857 CTCAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTGTGCACTGGGCGAGGAG 1916
Db 2002 CTCAAAAGCAGTCGTGAGGGCAGCTGGG--AGGACGTGTGCACTGGTGTGAGGAG 2058
QY 1917 GCGCCCTCCACCTCTACTACTGCTCCAGCTGGAGGCCAGTCCCTGCTACGTCCTTCAACGAG 1976
Db 2059 TCGCCCTCTCATCTACTACTGCTCCAGCTGGAGGCCGGGCGCTGCTATGTCCTCACCGAG 2118
QY 1977 CAGCTGGCGCGCTTTCGCCCTGTGGGAGAGGCCCTCAGCGTGGGTGCGCGCAGCGCCTC 2036
Db 2119 CAGCTAGCGCGCTTTCGCCCTGTGGGAGAGGCCCTCAGCGTGGGTGCGCGCAGCGCCTC 2178
QY 2037 AAGCTCTCTTGTGTGGCGCGTGGCTGCACCTCCCTCGAGTACACATCCCGGCTCTAC 2096
Db 2179 AGSCTCTCTTGTGGCCCTGTGGCTGTGTCCTCGAGTACAACTCCGAGTGTAC 2238
QY 2097 TGCCTGCATGACACCCAGATGCACTCAAGAGAGGTGTGAGTGGAGAGCAGCTGGGG 2156
Db 2239 TGCCTGCACGACACCCAGATGCTCTCAAGAGAGGTGTGAGTGGAGAGCAGCTGGGT 2298
QY 2157 GGACAGCTGATCAGGAGCCAGGGTCTGCACCTTCAGGACAGATTACCAACCTTGGCG 2216
Db 2299 GGACAGCTGATCAGGAGCCCGGTGTCGCACTTCAAAAGACAGTTACCAACCTTACCT 2358
QY 2217 CTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGTAAAGTCTCTGTGACGTACACG 2276
Db 2359 CTGTCCATCCAGACGTGCCAGCTCCCTGTGGAAGCAAGTCTCTGTGACGTACACG 2418
QY 2277 GAGATCCCTTTTATCAGATCTGGAATGGACGCGGTACTTGCATCTGCACCTTACC 2336
Db 2419 GAGATCCCTTTTATCAGATCTGGAATGGACGCGGTACTTGCATCTGCACCTTACC 2478
QY 2337 CTGGAGCGTGTAGCCCCAGCACTAGTGACCTGGCTGCAAGCTGTGGGTGTGGCAGGTG 2396
Db 2479 CTGGAGCGGTCAATGCCAGCACAGGACCTGGCCCTGCAAGGTGTGGGTGTGGCAGGTG 2538
QY 2397 GAGGCGGACGGGACAGCTTCAGCATCAACTTCAACATCAACAGGACACAAGGTTTGT 2456
Db 2539 GAGGAGATGGACAGACTTCAACATCACTTAACTACTAAGGACACAGAGTTTGT 2598
QY 2457 GAGCTGTGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCCCAAGTGCCTTC 2516
Db 2599 GAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTCCAGCCCTGGTGGGCCCAAGTGCCTTC 2658
QY 2517 AAGATCCCTTCTCATTCGCGAGAGATAATTTCCAGCTGGAACCCCTGTAGGCGG 2576
Db 2659 AAGATCCCTTCTCATTCGCGCAAGATATTTCCAGCTGGAACCCCTGTAGGCGG 2718
QY 2577 GGTGCGGACTGCGGACTCTGGCCAGAACTCCACTGACAGCACTCTAGCTTCTTT 2636
Db 2719 GCGCGGACTGCGGAACTCTAGCCAGAACTTCACTGGACAGCACTTCTAGCTTCTTT 2778
QY 2637 GCCTCGAAGCCCGACGACCCATGATCTCAACCTGTGGGAGCGCGGCACTTCCCC 2696
```

```
Db 2779 GCCTCCAGCCAGCCCTACAGCCATGATCTCAACCTATGGGAGGCGCGGCACTTCCCC 2838
QY 2697 AACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCACGACAGCGCTGGC 2756
Db 2839 AACGGCAACCTCAGCCAGCTGGCGCAGCTGTGGCGGACTGGGCCACGACAGTGTCTGGC 2898
QY 2757 CTCTTACAGCTGTCGAGGCTGAGTGTGAGGCGCGGCGAGGCCCG--ACACCTTACATCT 2814
Db 2899 CTCTTACAGCTGTCGAGGCGGAGTGTGAGACCCAGCCAGCGCGCTTACGCTTACATCT 2958
QY 2815 CACGAGCTTTGGCAACCCCAAGCAGGACAGGACCGGACAGGCGGCGCTTCCCCACAC 2874
Db 2959 CACGAGCTTTGGCAACCTTGGCAGGACAGGACAGGACAGGCGGCGCTTCCCCACAC 3018
QY 2875 GGGGAGA 2881
Db 3019 CCGGGA 3025

RESULT 6
BD057524
LOCUS Netrin receptors. 3014 bp DNA linear PAT 27-AUG-2002
DEFINITION BD057524
ACCESSION BD057524
VERSION BD057524.1 GI:22603130
KEYWORDS JP 2001505062-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 3014)
AUTHORS Lavigne,M.T., Leonardo,D.E., Hinck,L., Masu,M. and Masu,K.K.
TITLE Netrin receptors
JOURNAL Patent: JP 2001505062-A 1 17-APR-2001;
THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT PN JP 2001505062-A/1
PD 17-APR-2001
PF 19-FEB-1998 JP 1998536840
PR 19-FEB-1997 US 08/808982
PI MARC TESSIER LAVIGNE,DAVID E LEONARDO,LINDSAY HINCK,MASAYUKI
PI MASU,
PI KAZUKO KEINO MASU
PC C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
Feature 1..3014
Source /organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 77.6%; Score 2235.8; DB 6; Length 3014;
Best Local Similarity 88.5%; Pred. No. 0;
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

QY 87 ATGGCGCTCGGCGCGGCTGTGGCCAGCGCTTCCTGGGCATAGTCTCGCCGCTTGGCTC 146
Db 1 ATGGCGCTCGGCGCGGCTGTGGCCAGTCTCTCTGGGCATAGTCTCGCCGCTTGGCTT 60
QY 147 CGCGGCTCGGCTCGGCGCGGCTGTGGCCAGCGCTTCCTGGGCATAGTCTCGCTGGTGCACCG 206
Db 61 CGTGTTCGGGTGCGGCGCGGCTGTGGCCAGCGTTCCTGGGCATAGTCTCGCTGGTGCACCG 120
QY 207 GACCTGCTTCCCCTTCTCTGTGGAGCCCGGAGGATGTACATCGTCAAGACAGCA 266
Db 121 GACCTGCTTCCCCTTCTCTGTGGAGCTTCTGTAGAGCTTGTGATGTCAGAACAGCG 180
QY 267 GTGCTGTGTGTGCAAGGCGGTGCGCCCGCAGCAGATCTTCTTCAAGTCAACGGGAG 326
Db 181 GTGTTGTTGGTGTGCAAGGCTGTGCTGCCACCGCAGATCTTCTTCAAGTGCATGCGGNA 240
QY 327 TGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACGGGAGCAGTGGTGGCGG 386
```

241 TGGGTCGCGAGTTCGATCAGGTAATTGAACGAGCAGCAGCAGCGGATTCGCA 300  
387 ACCATGGAGTCCGCAATTAATGCTCAAGCAGCAGGTGAGAGGTGTTGGGCTGGAG 446  
301 ACCATGGAGTCCGATCAACGTATCGAGCAGCAGGTAGAGAAAGTGTGGGCTGGAG 360  
447 GAATACTGGTCCAGTGGGATGGAGTCTTCGGGCAACCAACGAAGTCAAGAGGCC 506  
361 GAATACTGGTCCAGTGGGATGGAGTCTTCGGGCAACCAACGAAGTCAAGAGGCC 420  
507 TACATCCGATAGCCAGATTGCCAGAACTTCGAGCAGGAGCGGTGGCCAAAGGAGGTG 566  
421 TACATCCGATAGCCATTTGGCCAGAACTTTGAGCAGGAGCACTGGCCAAAGGAGTG 480  
567 TCCTCGAGCAGGCGATCGTGTGCCCTGCGCTCCACCGGAGGCGATCCCTCCAGCGGAG 626  
481 TCACTGGAGCAAGCATGTTACTACCTTGTGCGCCGCCAGAGGAATCCCGCCAGCTGAG 540  
627 GTGGAGTGGTCCGGAAAGGAGTGGTGGACCGGTCCCTGGACCCCAATGTATATC 686  
541 GTGGAGTGGTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTATC 600  
687 ACGCGGAGCACAGCTGTGTGTCGACAGCGCGCTTGTGACACGGCCAACTACACC 746  
601 ACGCGGAGCACAGCTGTGTGTCGTCGAGCGCGCTTGTGACACGGCCAACTACACC 660  
747 TGGTGGCCAAAGAACATCGTGGCAGCTGCGCGCAGCGCTCCGCTGTGTCTATCTATC 806  
661 TGTGTGCCAAAGAACATCGTAGCCCGTGGCGAAGCACCTCTGCGCGGTCAITGTTAT 720  
807 GTGAACCGTGGTGGTTCGAGTGGACCGAGTGTGCTGCTGCGAGCGCAGCTGTGGCGC 866  
721 GTGAACCGTGGTGGTTCGAGTGGACCGAGTGTGCTGCTGCGAGCGCAGCTGTGGCGC 780  
867 GGCTGGCAGAAACGAGCGAGCTGACCAACCGCGCGCTCTCAACGGGGCGCTTTC 926  
781 GGCTGGCAGAAACGAGCGAGCTGACCAACCGCGCGCTCTCAACGGGGCGCTTTC 840  
927 TGTGAGGGGAGAAATGTCATAGCCGACCGCTGCTCTCTGTGTGTCTGTGTGGACGC 986  
841 TGTGAGGGGAGAAATGTCAGAAACAGCGCTGCGCCACTCTG---TGCCAGTGGATGG 897  
987 AGCTGGAGCCGTGGAGGAGTGTGCGCTGTGGCTGGACTGACCCACTGCGCGAGC 1046  
898 AGCTGGAGTTCGTGGAGTAAGTGTGACGCTGTGGCTTGACTGACCCACTGCGCGAGC 957  
1047 CGTCAGTGTCTGACCCAGACCCCGCAACGAGGGGAGGAGTGCACGGGCACCTGACCTG 1106  
958 CGCGAGTGTCTGACCCAGACCCCGCAATGGAGGTGAGGAGTGTGCGGGTGTCTGACCTG 1017  
1107 GACACCCGCAACTGTACAGTGAACCTGTGTGTACACAGTGTCTTGGCGCTGAGGAGTG 1166  
1018 GACACCCGCAACTGTACAGTGAACCTGTGCTGCACACCGCTTCTTGGCCCGAGGAGTG 1077  
1167 GCCTCTATGTGGCTCTATGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1226  
1078 GCTCTCTACATCGGCTTGTGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137  
1227 ATCTCTGTTATTGCCGGAAGAGGAGGCTGGACTCAGATGTGGCTGACTCGTCCATT 1286  
1138 GGACTCATTTACTGTGCGAAGAGGAGGCTGGACTCCGATGTGGCGACTCGTCCATC 1197  
1287 CTACCTCAGGCTTCAGCCCGTCAAGATCAAGCCCGAGCAAGCAGCAACCCCACTG 1346  
1198 CTCACCTCGGCTTCAGGCTGTGAGCATCAAGCCCGAGCAAGCAGCAACCCCACTG 1257  
1347 CTCACCTCAGGCTTCAGCCCGTCAAGATCAAGCCCGAGCAAGCAGCAACCCCACTG 1403  
1258 CTCACCTCAGGCTTCAGGCTGTGAGCATCAAGCCCGAGCAAGCAGCAACCCCACTG 1317  
1404 CGCAGGATGGCCCGAGCCCAAGTTCAGTCTCAACATGGGACCTGCTCAGGCCCTG 1463

1318 AGCGAGATGAGCCAGCCAGCCCAAGTTCAGCTCTCTAATGTTCCAGCTCACCTGTCCAGCCACTG 1377  
1464 GGTGGGGCGGCCACACACTGACACAGCTCTCCAGCTCTGAGCCCGAGGATTCGTC 1523  
1378 GCGAGTGGCGCCCATAGTTGACCAAGTTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437  
1524 TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCCGAGGCAACCAACATGACC 1583  
1438 TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCCGAGGCAACCAATGACC 1497  
1584 TATGGACCTTCAATCTCTCCGCGGCGCTGATGATCCCTAATACAGGATACGCTC 1643  
1498 TACCGACCTTCAATCTCTCCGCGGCGCTGATGATCCCTAATACAGGATACGCTC 1557  
1644 CTATCCCGCCAGATCCCATACCCGAGGAGATCTATGAGATCTACCTTCACTGAC 1703  
1558 CTATACCCCGGATCCCATCCCGAGGAAAGATCTACGAGATCTACCTTCACTGAC 1617  
1704 AAGCCGGAAGACCTGAGGTTGCCCTAGTGGTGTGACACCCCTGTGAGTCCCATCGTT 1763  
1618 AAGCCAGAGACGTGAGGTTGCCCTAGTGGTGTGACACCCCTGTGAGTCCCATCGTT 1677  
1764 AGCTGTGACCCCTCGCGTCTGCTCAACCGGCGAGTCACTGCTGATGAGCACTGT 1823  
1678 AGCTGTGGGCGCCAGAGTCTGCTCACCCGCGCAGTCACTCTGCAATGAGCACTGT 1737  
1824 GGGAGCCAGCCCTGACAGCTGGAGCTGCGCTCAAAAGCAGTCTGCGAGGGCAGC 1883  
1738 GAGAGCCAGCCCTGACAGCTGGAGTCTGCGCTCAAAAGCAGTCTGCGAGGGCAGT 1797  
1884 TGGAGCAGGATGTGTGCACTTGGGCGAGGCGCCCTCCACCTCTACTACTGTCAG 1943  
1798 TGGG---AGGATGTGTGCACTTGGTGGAGGAGTCACTTCCACCTCTACTACTGTCAG 1854  
1944 CTGGAGCCAGTGTGCTGCTGCTTCCAGCAGAGTGTGGCGCTTTCCTGCTGGTGGGA 2003  
1855 CTGGAGCCGCGGCTGTATGCTTTCAGCAGCAGCTGGCGCGCTTTCCTGCTGGTGGGA 1914  
2004 GAGCCCTCAGCTGGCTGCGCCAGCGCTCAAGCTGCTTCTGTTTGGCGCGTGGCC 2063  
1915 GAGCCCTCAGCTGGCTGCGCCAGCGCTCAAGCGCTCTTCTGTTTGGCGCGTGGCC 1974  
2064 TGCACCTTCCCTCGAGTACAACTCCCGGCTTACTGCTGTGATGACACCCAGTGCATC 2123  
1975 TGTACGTCCTTGTGATACACATCCGAGTGTACTGCTTCAACGACACCCAGCGCTC 2034  
2124 AAGAGTGTGTGAGTGGAGAGAGCTGGGGGAGCAGCTGATCCAGAGGACCGGGTC 2193  
2035 AAGAGTGTGTGAGTGGAGAGAGCTGAGTGGAGCAGCTGATCCAGGAGCTTCGCGTC 2094  
2184 CTGCACCTCAAGGACAGTACCAACCTGCGCTATCCATCCAGATGTGCCAGCTCC 2243  
2095 CTGCACCTCAAGGACAGTACCAACCTGAGTGTCTTCCATCCAGAGTGTGCCAGCTCC 2154  
2244 CTGTGGAAGAGTAAAGCTTGTGTCAGTACAGGAGATCCCTTTTATCATCTGGAAT 2303  
2155 CTGTGGAAGAGTAAAGCTTGTGTCAGTACAGGAGATCCCTTTTATCCACATCTGGAAC 2214  
2304 GGCACGAGCGGTACTTGGCTGCACTTCACTTCACTGAGCGTGTGAGCCCGCAGCTAGT 2363  
2215 GGCACCGAGCAGTATGCACTTCACTTCACTGAGCGCATCAACCGCAGCAGCAGC 2274  
2364 GACCTGCGCTGCAAGCTGTGGTGTGAGGCGGCGAGGCGAGAGCTTCAGCATC 2423  
2275 GACCTGCGCTGCAAGTGTGGTGTGAGGAGTGTGAGGAGAGTGGCAGAGCTTCAACATC 2334  
2424 AACTTCAACATCAACAGGACAAAGGTTGTGCTGAGTGTGCTGCTGAGAGTGAAGCG 2483  
2335 AACTTCAACATCAACAGGACAAAGGTTGTGCTGAAATTTGCTGCTGAGAGTGAAGGG 2394  
2484 GGGTTCACCGCTGTTGGCGCCAGTCCCTTCAAGATCCCTTCTTCTTCTTCTTCTTCT 2543  
2395 GGGTTCACCGCTGTTGGCGCCAGTCCCTTCAAGATCCCTTCTTCTTCTTCTTCTTCT 2454



1464 GTTGGGGGCGGCAACACATGACACAGCTCTCCACCTCTGAGGCCGAGGATTCGTC 1523  
1378 GGGAGTGGCGGCGCATACGTTGACACACAGCTACACCCACCTCTGAGGCTGAGGATTCGTC 1437  
1524 TCCGCGCTCTCCACCCAGACTACTTCCGCTCCCTGCGCGGCGGACACACAGCAATGACC 1583  
1438 TCCGCGCTCTCCACCCAACTACTTCCGCTCCCTGCGCGGCGGACACAGCAATGAGCC 1497  
1584 TATGGACCTTCAACTCTCTCGGGGCGGCTGATGATCCCTAATACAGGATATCAGCCTC 1643  
1498 TACGGACCTTCAACTCTCTCGGGGCGGCTGATGATCCCTAATACAGGATATCAGCCTC 1557  
1644 CTATCCCCCAGATGACATACCCGAGGAGAGATCTATGAGATCTACCTCAGCTGCAC 1703  
1558 CTATACCCCGGATGACATCCCGGAGGAGAGATCTACGAGATCTACCTCAGCTGCAC 1617  
1704 AAGCGGAGAGAGCTGAGTTGCCCTAGCTGGCTGTGACAGCCCTGCTGAGTCCCATCGTT 1763  
1618 AAGCGGAGAGAGCTGAGTTGCCCTAGCTGGCTGTGACAGCCCTGCTGAGTCCCATCGTT 1677  
1764 AGCTGTGGACCCCTGCTGCTGCTCACCGGCGAGTCACTCTGGCTATGGAACCACTGT 1823  
1678 AGCTGTGGGCCCCAGAGTCTGCTCACCGGCGGAGTCACTCTGGCTATGGAACCACTGT 1737  
1824 GGGAGCCCGGCTGACAGCTGGAGCTGCGCTCAAAAAGAGTCTGCGAGGGGAGC 1883  
1738 GAGAGCCCGGCTGACAGCTGGAGTCTGCGCTCAAAAAGAGTCTGCGAGGGGAGT 1797  
1884 TGGAGAGAGATGCTGACCTGGGCGGAGGAGCGCTCCCACTCTACTACTGCGAG 1943  
1798 TGGG--AGGATGCTGACCTTGGTGAGAGTCACTCTCCACCTCTACTACTGCGAG 1854  
1944 CTGAGGCGCAGTGTGCTGCTGCTTCCAGGAGAGCTGGGCGGCTTGGCTGGTGGGA 2003  
1855 CTGAGGCGGCGGCTGCTGCTTCTTCCAGGAGAGCTGGGCGGCTTGGCTGGTGGGA 1914  
2004 GAGCCCTCAGCTGTGCTGCGGCGGAGGCTCAAGGCTCAAGCTGCTTGGTGGGCGGCTG 2063  
1915 GAGCCCTCAGCTGTGCTGCGGCGGAGGCTCAAGGCTCAAGCTGCTTGGTGGGCGGCTG 1974  
2064 TGACCTCTCCGAGTCAACATCGGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2123  
1975 TGTACGCTCTTGGTCAACATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034  
2124 AAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2183  
2035 AAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2094  
2184 CTGCACTTCAAGGAGAGTACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2243  
2095 CTGCACTTCAAGGAGAGTACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2154  
2244 CTGTGGAAGAGTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2303  
2155 CTGTGGAAGAGTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2214  
2304 GGCACCGAGCGGTACTTGCATGCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2363  
2215 GGCACCGAGCGGTACTTGCATGCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274  
2364 GACCTGGCTGCAAGCTGTGGTGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2423  
2275 GACCTGGCTGCAAGCTGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2334  
2424 AACTTCAACATCAACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2483  
2335 AACTTCAACATCAACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2394  
2484 GGGGTCCAGCGCTGTGGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2543  
2395 GGGGTCCAGCGCTGTGGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2454

2544 ATAAATTTTCAGCTGTGACCCACCTCTAGCGGGGTGCGGAGTCTGCGGAGTCTGCGGCCAG 2603  
2455 ATCAATCGCCAGTCTGACCCACCTCTGACCGGGGGCGGAGTCTGAGAGAACTCTAGCCAG 2514  
2604 AAATCTCACTGTGAGAGCCATCTCAGCTCTTTTGTCTTCAAGCCAGCCAGCCAGCCATG 2663  
2515 AAATCTCACTGTGAGAGCCATCTCAGCTCTTTTGTCTTCAAGCCAGCCAGCCATG 2574  
2664 ATCTCACTGTGAGAGCCATCTCAGCTCTTTTGTCTTCAAGCCAGCCAGCCATG 2723  
2575 ATCTCACTGTGAGAGCCATCTCAGCTCTTTTGTCTTCAAGCCAGCCAGCCATG 2634  
2724 GCACTGTGCTGAGTCTGCGGCGGAGCCAGCTGCTCTTCAAGTCTGCGAGGCTGAGTGC 2783  
2635 GCTGTGCGGAGTCTGCGGCGGAGCCAGCTGCTCTTCAAGTCTGCGAGGCTGAGTGC 2694  
2784 TGA 2786  
2695 TGA 2697

## RESULT 8

RN087305  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 2697)  
Leonardo, E.D., Hincik, L., Masu, M., Keino-Masu, K., Ackerman, S.L. and  
Tessier-Lavigne, M.  
Vertebrate homologues of C. elegans UNC-5 are candidate netrin  
receptors  
Nature 386 (6627), 833-838 (1997)  
97271897  
9126742  
2 (bases 1 to 2697)  
Leonardo, E.D., Hincik, L., Masu, M., Keino-Masu, K. and  
Tessier-Lavigne, M.  
Direct Submission  
Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San  
Francisco, CA 94143-0452, USA  
FEATURES  
Location/Qualifiers  
1..2697  
/organism="Rattus norvegicus"  
/mol\_type="mrna"  
/db\_xref="taxon:10116"  
/tissue\_type="brain and ventral spinal cord"  
/dev\_stage="18 day embryo and 13 day embryo"  
1..2697  
/codon\_start=1  
/product="transmembrane receptor UNC5H1"  
/protein\_id="AAB57678.1"  
/db\_xref="GI:2055392"

## CDS

/translation="MAVRPGLWPLVLLGIIVLAAMLRGSGAQSATVANPVCANPDLLP  
HFLPEPEVIVKNNPVLLCKAVPATQIFKNGEWVRQVDHVIERTSDSSGLPTM  
SLRINVRQVQKVEFLEWYCCQVWSSGTTKSKAYIRIAYLRNFEPIAKEV  
ELEGIVLPKPPPEGIIPAEVWLRNEDLVDPNIVYITREHSLVVRQARLADTAN  
YTCVAKNIVARRRSTSAIVVYNGWSTWESVCSGCRGWRKRSCTNPAPLN  
GGACCEQNVQKTACATLCPVDGWSWSKWSACGLDCTHWRSECSDPAPNGGEEC  
RGALDTRNCTSDLCITASCPEVALYIGLVAVAVCLFLLLLGLIYCRKKEGLDS  
DVADSSILTGFPVSIKPSKADNPILLTIQPDLSSTTTTQGSICSRQDGSPKFLQ  
SNHLLSPGSGRTHLHSSSEADPFVSLSTYSPRSLPRGTNNMAYGTFFLGG  
RLMPTNGTISLLIPDAIPIKYEIVITLHKPEDVRLPLAGCQTLLSPVSCGPPGV  
LLTPVILAMDHCCEPSPDSRLKQSCGSEWEDVLHGEESPSHLYYQQLPAGAC  
YVTFEQLGRPALVGEALSVAKTRLLPAPVACTSLEYNIRVYCLADTHDLKEVV  
QLEKQLGQLLQEPFVLFKDSYHNRLSLIHDPSSLWKSLLYSQEIPEVHTWNGT  
QQYLHCTFTLERINASTDSLACKVWVQVGGGGSFNINFTKDTFAELLLESEG









GTQVYLHCTFTLERNVASTSLACKVWVQVGDGQSFNINFINITKDRFAEMLALES  
EGYPAIVGPAFKIPILIRQKIITSLDPPGSRGADWRITLAKHLHLSLHFFASKPS  
TAMILNLWEARHPNGNLGLAAAVAGLQPDAGLFTVSEAC"  
714. .926  
/note="IG; Region: Immunoglobulin"  
/db\_xref="CDD:smart00409"  
984. .1130  
/note="TSPI; Region: Thrombospondin type 1 repeats"  
/db\_xref="CDD:smart00209"  
1566. .1877  
/note="ZU5; Region: Domain present in ZO-1 and Unc5-like  
netrin receptors"  
/db\_xref="CDD:smart00219"  
2496. .2684  
/note="DEATH; Region: DEATH domain, found in proteins  
involved in cell death (apoptosis). Alpha-helical domain  
present in a variety of proteins with apoptotic functions.  
Some (but not all) of these domains form homotypic and  
heterotypic dimers"  
/db\_xref="CDD:smart00005"

Query Match 70.1%; Score 2020.6; DB 10; Length 3844;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 2429; Conservative 0; Mismatches 274; Indels 184; Gaps 6;

QY 1 AGCTGGGGCTCCGGGCTCAGCGCT-AAAGCGCGCTCCCGCGCGGGGCCCGCGCCC 59  
Db 169 AGGTGGGGCTCTCGGCTGAGCGCTGACGGCTCTCTCCGCGCGGGGCCCTAGTCC 228  
QY 60 GGGCGGCGCCCTGCGCGCGCGGCGCATGGCGGTCCGGCGCGGCGCTGTGGCGCGCTC 119  
Db 229 AGCCCGCTCGT-----CGCGCGCGGCGCATGGCGGTCCGGCGCGGCGCTGTGGCGCGCTC 284  
QY 120 CTGGGATAGTCTCGCGCTCGCGCTCGCGGCTCGGGTCCCGAGAGAGTGCCACCGTG 179  
Db 285 CTGGGATAGTCTCGCGCTCGCGCTCGCGGCTCGGGTCCCGAGAGAGTGCCACAGTG 344  
QY 180 GCCAACCCAGTGGTGTGCAACCGCGACCTGTCTCCCGACCTTCTGTGGAGCGCGAG 239  
Db 345 GCCAACCCAGTGGTGTGCAACCGCGACCTGTCTCCCGACCTTCTGTGGAGCGCGAG 404  
QY 240 GATGTGATCATCTGAAGAACAGCCAGTGTCTGTGTGCAAGGCGGTGCGCGCCACG 299  
Db 405 GAGTGTATCATTTCAAGAACAGCCCGTGTCTGTGTGCAAGGCGGTGCGCGCCACC 464  
QY 300 CAGATCTTCTCAAGTGCAACCGGGAGTGGTTCGCGCAGGTGACCAAGTATCGAGCGC 359  
Db 465 CAGATCTTCTCAAGTGCAACCGGGAGTGGTTCGCGCAGGTGACCAAGTATCGAGCGC 524  
QY 360 AGCAGACGGGAGTGTGTGAGCGGACCATGGAGGTCCGCATTAAATGTCTCAAGGCAG 419  
Db 525 AGCACTGACGGCAGCAGCGGATTGCCAACATGGAGGTCCGATCAACGTATCAAGGCAG 584  
QY 420 CAGGTGCAAGAGTGTTCGGGTGGAAGAACTGTGTGCAAGTGTGCGTGGCATGGAGTCC 479  
Db 585 CAGGTGCAAGAAAGTGTTCGGGTGGAAGAACTGTGTGCAAGTGTGCGTGGCATGGAGTCC 644  
QY 480 TCGGCGACCAACAGAGTCAAGAGGCTACATCCGATAGCAGATGGCGAAGAACTTC 539  
Db 645 TCAGNACCACCAAGCAGAGGCTACATCCGGATTGCGCTATTTCGGCAAGAACTTT 704  
QY 540 GAGCAGGAGCGGTGGCCAAAGAGGTGTCTCTGGAGCAGGCGATCGTGTGCGCTGCCGT 599  
Db 705 GAGCAGGAGCGGTGGCCAAAGAGTGTCTCTGGAGCAGGCGATTTGTACTTGTGCGC 764  
QY 600 CCACGGAGGGCATCTCTCCAGCGAGGTGAGTGGCTCCGGAAACGAGGACCTGTGTGAC 659  
Db 765 CCCCCGGAAGGAATCCCCCAGCTGAGGTGGAGTGGCTCCGAAATGAGGACCTGTGTGAC 824  
QY 660 CGGTCTCTGACCCCAATGTATACATCAGCGGGAGCAGCGCTGTGTGTGCGACAGGCC 719  
Db 825 CCTCTCCTCGACCCCAATGTATACATCAGCGGGAGCAGCGCTGTGTGTGCGGAGGCC 884

ORIGIN

QY 720 GCGCTTGTGTGACACGGCCAACTACACTGCTGTCGCAAGAAACATCGTGGCAGCTGCCGC 779  
Db 885 GCGCTGCGCAGCACTCCCAACTACACTGCTGTCGCAAGAAACATCGTGGCCGTCGCGCA 944  
QY 780 AGCGCTTCCGCTGCTGTCATCTACTGTAACGGTGGTGGTGTGACGTGGACCGAGTGG 839  
Db 945 AGCGCTTCTGGCGCGTCA----- 963  
QY 840 TCGCTCTGACGCGCAGCTGTGGCGCGGCTGGCAGAAACGAGCGGAGCTGCACCAAC 899  
Db 964 ----- 963  
QY 900 CCGCGCGCTCTCAACGGGGGGCTTTCTGTGAGGGGAGAAATGTCATGACCGCACCGTC 959  
Db 964 ----- 963  
QY 960 TCCTCTCTGCTGTGCTCTGTGGACGGCAGCTGGAGCCGCTGGAGCAAGTGGTCCGCCCTGT 1019  
Db 964 -----TTGTTTATGTGGATGGAGCTGGAGCCCATGGAGTAAGTGGTCAAGCTGC 1013  
QY 1020 GGGCTGACTGCAACCCCACTGCGGAGCCGTGAGTGTCTGACCCAGCACCCCGCAACGGA 1079  
Db 1014 GGGCTTGACTGCAACCCCACTGCGGAGCCGGAGTGTCTCGACCCAGCGCCCGCAACGGA 1073  
QY 1080 GGGGAGAGTGCAGGGCACTGACTGGACACCCGCAACTGTACCAAGTACCTTGTGTGA 1139  
Db 1074 GGTGAGAGTGCAGGGGTGTGACTGGACACCCGCAACTGTACCAAGTACCTTGTGTGA 1133  
QY 1140 CACAGTGTCTTCCGCTGAGGAGCTGGCCCTCTATGTGGGCTCATCGCGCTGTGGCGTC 1199  
Db 1134 CACACCTTTCGCGCCCGGAGAGCTGGCTCTCATATGGGCTCTGCGCGTGTGGCGTG 1193  
QY 1200 TGCTGTCTGTGTGTGTCTCTCATCTCTGTTTATTCGCGGAGAGAGGAGGCTG 1259  
Db 1194 TGCTCATCTGTGTGTGTGTCTCTCATCTCTGCTCTCATCTACTGCGCAGAGGAGACTG 1253  
QY 1260 GACTGAGTGTGGTGACTGCTCATCTCTCAGGCTTCCAGGCTTCCAGGCTGTGAGCACTCAAG 1319  
Db 1254 GACTGAGAGTGGTGACTCATCTCTTACCTCAGGCTTCCAGGCTGTGAGCACTCAAG 1313  
QY 1320 CCAGCAAGAGCAGCAACCCCATCTGCTCACCATCCAGCGGACCTCAGCACTCACCC 1376  
Db 1314 CCAGCAAGAGCAGCAACTCCCATCTGCTCACCATCCAAACCGGAGCTCAGCACTCACCC 1373  
QY 1377 ACCACCTACAGGGCAGTCTCTGTCCCGCGCAGATGGGCGGAGAGGAGGAGGCTG 1436  
Db 1374 ACCACCTACAGGGCAGCTGTGTCCCGCGCAGATGGGCGGAGAGGAGGAGGCTG 1433  
QY 1437 ACCAATGGGACCTGTGTGAGCGGCTGTGGTGGCGGCGGCGGAGGAGGAGGAGGAG 1496  
Db 1434 TCTAATGGTCACTGTGTGAGCGGCTGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAG 1493  
QY 1497 CCGACCTGTGAGCGGAGGAGTGTCTCCCGCTCTCCAGCGGAGGAGGAGGAGGAGGAG 1556  
Db 1494 CCGACCTGTGAGGCTGAGGAGTGTCTCCCGCTCTCCAGCGGAGGAGGAGGAGGAGGAG 1553  
QY 1557 CTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1616  
Db 1554 CTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1613  
QY 1617 ATGATCCCTAATACAGGATCAGCTCTCATCTCCCGGAGGAGGAGGAGGAGGAGGAGGAG 1676  
Db 1614 ATGATCCCTAATACAGGATCAGCTCTCATCTCCCGGAGGAGGAGGAGGAGGAGGAGGAG 1673  
QY 1677 ATCTATGAGTCTACTCAGCTGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1736  
Db 1674 ATCTAGAGATCTACTCAGCTGACAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1733  
QY 1737 TGTGAGACCTGCTGAGTCCCATCTGATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1796  
Db 1734 TGTGAGACCTGCTGAGTCCCATCTGATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1793

QY	1797	CCAGTCATCTCGCTATGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGCG	1856
Db	1794	CCAGTCATCTCTTGGCATGACCACTGTGGGAGGCCAGCTCCGACAGCTGGAGCTGGCG	1853
QY	1857	CTCAAAAGCAGTCGTGGCAGGSCAGCTGGAGCAGGATGTGCTGCACCTGGCGGAGGAG	1916
Db	1854	CTCAAAAGCAGTCGTGGGAGCAGCTGGG---AGGACGTGCTGCACCTGTGTGAGGAG	1910
QY	1917	GGCCCTCCCACTCTACTACTCCAGCTGGAGGCCAGTGCCTGTACCTTTTACCAGAG	1976
Db	1911	TCGCCCTCTCATCTACTACTGCCAGCTGGAGGCCCTGTATGTCTTTTACCAGAG	1970
QY	1977	CAGCTGGCGGCTTTGGCCCTGTGGAGAGGCCCTCAGCGTGGCTGCCGACGCGCTC	2036
Db	1971	CAGCTAGGCGGCTTTGCCCTGTGGAGAGGCCCTCAGCGTGGCTGCCCAAGCGGCTC	2030
QY	2037	AAGCTGCTCTCTTTGGCGCGGTGGCTGCACCTCCTCAGTACACATCCCGGTCTAC	2096
Db	2031	AGGCTGCTCTCTTTGGCCCTGTGGCTGTACGTCTCCTCAGTACACATCCGAGTGAC	2090
QY	2097	TGCTGATGACACCCACGATGACCTCAAGAGGTGTGAGTGGAGAGAGCTGGGG	2156
Db	2091	TGCTGACGACACCCACGATGCTCTCAAGAGGTGTGAGTGGAGAGAGCTGGGT	2150
QY	2157	GGACGCTGATCGAGGAGCCAGGCTCTGCACCTTCAAGACAGTTTACCAACCTCGCG	2216
Db	2151	GGACGCTGATCGAGGAGCCCGTCTCTGCACCTTCAAGACAGTTTACCAACCTACGT	2210
QY	2217	CTATCATTCACGATGTGCCAGCTCCTCTGGAAGAGTAACTCTTTCAGTACACAG	2276
Db	2211	CTGTCATTCACGAGTGTGCCAGCTCCTCTGGAAGAGTAACTCTTTCAGTACACAG	2270
QY	2277	GAGATCCCTTTTATCATCTGGAATGGCAGCGGTTACTTGACTGCACCTTCACC	2336
Db	2271	GAGATCCCTTTTATCATCTGGAATGGCAGCTTACTGCACCTTCACCTTCACC	2330
QY	2337	CTGGAGGCTGTGACCCACGACTAGTGTGCTGCTGCAAGCTGTGGGTGTCAGAGTG	2396
Db	2331	CTGGAGGCTGTGAAATGACACGAGCTGCTGCTGCAAGCTGTGGGTGTCAGAGTG	2390
QY	2397	GAGGCGCAGCGGAGAGCTTCAGCATCAACTTCAACATCAACAGGACACAAAGTTTGT	2456
Db	2391	GAGGAGATGAGCAGAGCTTCACATCAACTTCAACATCAACAGGACACAAAGTTTGT	2450
QY	2457	GAGTGTGCTCTGAGAGTGAAGCGGGGTCCAGCCTGTGTGGGCCCCAGTGCCTTC	2516
Db	2451	GAAATGTGCTCTGAGAGTGAAGGGGGGTCCAGCCTGTGTGGGCCCCAGTGCCTTC	2510
QY	2517	AAGATCCCTTCTCATTCGCGCAGAGATATTTCCAGCCTGGACCCACCTGTAGCGGG	2576
Db	2511	AAGATCCCTTCTCATTCGCGCAGAGATATTTCCAGCCTGGACCCACCTGTAGCGGG	2570
QY	2577	GGTCCGACTGGCGGACTCTGGCCCGAGAACTCCACCTGGACAGCCATCTCAGCTTCTTT	2636
Db	2571	GGCGCCGACTGGCGAACTCTAGCCCGAGAACTTCACCTGGACAGCCATCTTAGCTTCTTT	2630
QY	2637	GCCTCCAAAGCCAGCCCGCAGCCATGATCTCACTGTGGAGGCGCGCATTCCCC	2696
Db	2631	GCCTCCAAAGCCAGCCCGCAGCCATGATCTCACTGTGGAGGCGCGCATTCCCC	2690
QY	2697	AAGGCAACCTCAGCAGCTGGCTGACAGCTGGCTGAGCTGGGCGACCCAGAGAGCTGGC	2756
Db	2691	AAGGCAACCTCAGCAGCTGGCTGAGCTGGCTGAGCTGGGCGACCCAGAGAGCTGGC	2750
QY	2757	CTCTTCACAGTGTGGAGGCTGAGTGTGAGGCGCGGCCAGCCCGACAC - CCTACACTCT	2814
Db	2751	CTCTTCACGCTGAGAGGCGGAGTGTGAGCAGCAGCGCGGTACAGCTCTACACTCT	2810
QY	2815	CACGAGCTTTGGCACCACACAGGACAGCAGGAGCCGAGAGGCGGCTTCCCCACAC	2874
Db	2811	CACGAGCTTTGGCACCCTCCAGGGGACAGGCAAAAGCCAGAGGCGGCTTACCCCAAC	2870
QY	2875	GGGGGAGA	2881

FEATURES

source

1. 2688  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

Db	2871	CCGGGGA	2877
RESULT 10	BC009333	2688 bp	linear
LOCUS	BC009333	2688 bp	linear
DEFINITION	Homo sapiens cDNA clone IMAGE:4126760, partial cds.		
ACCESSION	BC009333		
VERSION	BC009333.2	GI:40226527	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2688) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalobio,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,K.Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
REFERENCE	2 (bases 1 to 2688) Strausberg,R.		
AUTHORS	Direct Submission		
TITLE	Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
JOURNAL	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
REMARK	On Dec 19, 2003 this sequence version replaced gi:14424611.		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-x@mail.nih.gov">cgabs-x@mail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc.mgc@nih.gov">nisc.mgc@nih.gov</a> Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripoop,S., Thomas,P.J., Touchman,J.W., Tsurgouon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,I., Young,A., Zhang,L.-H. and Green,E.D.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAL Plate: 26 Row: g Column: 22.		
source	Location/Qualifiers 1. 2688 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"		

<1. .1627

```
/codon_start=2
```

/protein\_id="A

```
/translation="")
```

GGGRHTLHHSSPTSI

MDHCGEPSPD SWSL

LIQEPRVLHFKDSYH

AFKIPFLIRQKIS

413. 724

netrin receptors, re-

1343 1594

involved in cell death; **POLE**==

```

/ab_xref="CD:3:

```

670

purity 99.6%; PI

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

[illegible]

\_\_\_\_\_

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

100

01234567891011121314151617181920212223242526272829303132333435363738394041424344454647484950515253545556575859606162636465666768697071727374757677787980818283848586878889909192939495969798991001011021031041051061071081091101111121131141151161171181191201211221231241251261271281291301311321331341351361371381391401411421431441451461471481491501511521531541551561571581591601611621631641651661671681691701711721731741751761771781791801811821831841851861871881891901911921931941951961971981992002012022032042052062072082092102112122132142152162172182192202212222232242252262272282292302312322332342352362372382392402412422432442452462472482492502512522532542552562572582592602612622632642652662672682692702712722732742752762772782792802812822832842852862872882892902912922932942952962972982993003013023033043053063073083093103113123133143153163173183193203213223233243253263273283293303313323333343353363373383393403413423433443453463473483493503513523533543553563573583593603613623633643653663673683693703713723733743753763773783793803813823833843853863873883893903913923933943953963973983994004014024034044054064074084094104114124134144154164174184194204214224234244254264274284294304314324334344354364374384394404414424434444454464474484494504514524534544554564574584594604614624634644654664674684694704714724734744754764774784794804814824834844854864874884894904914924934944954964974984995005015025035045055065075085095105115125135145155165175185195205215225235245255265275285295305315325335345355365375385395405415425435445455465475485495505515525535545555565575585595605615625635645655665675685695705715725735745755765775785795805815825835845855865875885895905915925935945955965975985996006016026036046056066076086096106116126136146156166176186196206216226236246256266276286296306316326336346356366376386396406416426436446456466476486496506516526536546556566576586596606616626636646656666676686696706716726736746756766776786796806816826836846856866876886896906916926936946956966976986997007017027037047057067077087097107117127137147157167177187197207217227237247257267277287297307317327337347357367377387397407417427437447457467477487497507517527537547557567577587597607617627637647657667677687697707717727737747757767777787797807817827837847857867877887897907917927937947957967977987998008018028038048058068078088098108118128138148158168178188198208218228238248258268278288298308318328338348358368378388398408418428438448458468478488498508518528538548558568578588598608618628638648658668678688698708718728738748758768778788798808818828838848858868878888898908918928938948958968978988999009019029039049059069079089099109119129139149159169179189199209219229239249259269279289299309319329339349359369379389399409419429439449459469479489499509519529539549559569579589599609619629639649659669679689699709719729739749759769779789799809819829839849859869879889899909919929939949959969979989991000100110021003100410051006100710081009101010111012101310141015101610171018101910201021102210231024102510261027102810291030103110321033103410351036103710381039104010411042104310441045104610471048104910501051105210531054105510561057105810591060106110621063106410651066106710681069107010711072107310741075107610771078107910801081108210831084108510861087108810891090109110921093109410951096109710981099110011011102110311041105110611071108110911101111111211131114111511161117111811191120112111221123112411251126112711281129113011311132113311341135113611371138113911401141114211431144114511461147114811491150115111521153115411551156115711581159116011611162116311641165116611671168116911701171117211731174117511761177117811791180118111821183118411851186118711881189119011911192119311941195119611971198119912001201120212031204120512061207120812091210121112121213121412151216121712181219122012211222122312241225122612271228122912301231123212331234123512361237123812391240124112421243124412451246124712481249125012511252125312541255125612571258125912601261126212631264126512661267126812691270127112721273127412751276127712781279128012811282128312841285128612871288128912901291129212931294129512961297129812991300

GGGAGGATGGGCTT

**C  
C  
C  
C  
C  
C  
C  
C  
E  
C  
C**

1. **Introduction**  
 2. **Background**  
 3. **Methodology**  
 4. **Results**  
 5. **Discussion**  
 6. **Conclusion**  
 7. **References**  
 8. **Appendix**  
 9. **Index**  
 10. **Table of Contents**  
 11. **Figure 1**  
 12. **Figure 2**  
 13. **Figure 3**  
 14. **Figure 4**  
 15. **Figure 5**  
 16. **Figure 6**  
 17. **Figure 7**  
 18. **Figure 8**  
 19. **Figure 9**  
 20. **Figure 10**  
 21. **Figure 11**  
 22. **Figure 12**  
 23. **Figure 13**  
 24. **Figure 14**  
 25. **Figure 15**  
 26. **Figure 16**  
 27. **Figure 17**  
 28. **Figure 18**  
 29. **Figure 19**  
 30. **Figure 20**  
 31. **Figure 21**  
 32. **Figure 22**  
 33. **Figure 23**  
 34. **Figure 24**  
 35. **Figure 25**  
 36. **Figure 26**  
 37. **Figure 27**  
 38. **Figure 28**  
 39. **Figure 29**  
 40. **Figure 30**  
 41. **Figure 31**  
 42. **Figure 32**  
 43. **Figure 33**  
 44. **Figure 34**  
 45. **Figure 35**  
 46. **Figure 36**  
 47. **Figure 37**  
 48. **Figure 38**  
 49. **Figure 39**  
 50. **Figure 40**  
 51. **Figure 41**  
 52. **Figure 42**  
 53. **Figure 43**  
 54. **Figure 44**  
 55. **Figure 45**  
 56. **Figure 46**  
 57. **Figure 47**  
 58. **Figure 48**  
 59. **Figure 49**  
 60. **Figure 50**  
 61. **Figure 51**  
 62. **Figure 52**  
 63. **Figure 53**  
 64. **Figure 54**  
 65. **Figure 55**  
 66. **Figure 56**  
 67. **Figure 57**  
 68. **Figure 58**  
 69. **Figure 59**  
 70. **Figure 60**  
 71. **Figure 61**  
 72. **Figure 62**  
 73. **Figure 63**  
 74. **Figure 64**  
 75. **Figure 65**  
 76. **Figure 66**  
 77. **Figure 67**  
 78. **Figure 68**  
 79. **Figure 69**  
 80. **Figure 70**  
 81. **Figure 71**  
 82. **Figure 72**  
 83. **Figure 73**  
 84. **Figure 74**  
 85. **Figure 75**  
 86. **Figure 76**  
 87. **Figure 77**  
 88. **Figure 78**  
 89. **Figure 79**  
 90. **Figure 80**  
 91. **Figure 81**  
 92. **Figure 82**  
 93. **Figure 83**  
 94. **Figure 84**  
 95. **Figure 85**  
 96. **Figure 86**  
 97. **Figure 87**  
 98. **Figure 88**  
 99. **Figure 89**  
 100. **Figure 90**  
 101. **Figure 91**  
 102. **Figure 92**  
 103. **Figure 93**  
 104. **Figure 94**  
 105. **Figure 95**  
 106. **Figure 96**  
 107. **Figure 97**  
 108. **Figure 98**  
 109. **Figure 99**  
 110. **Figure 100**  
 111. **Figure 101**  
 112. **Figure 102**  
 113. **Figure 103**  
 114. **Figure 104**  
 115. **Figure 105**  
 116. **Figure 106**  
 117. **Figure 107**  
 118. **Figure 108**  
 119. **Figure 109**  
 120. **Figure 110**  
 121. **Figure 111**  
 122. **Figure 112**  
 123. **Figure 113**  
 124. **Figure 114**  
 125. **Figure 115**  
 126. **Figure 116**  
 127. **Figure 117**  
 128. **Figure 118**  
 129. **Figure 119**  
 130. **Figure 120**  
 131. **Figure 121**  
 132. **Figure 122**  
 133. **Figure 123**  
 134. **Figure 124**  
 135. **Figure 125**  
 136. **Figure 126**  
 137. **Figure 127**  
 138. **Figure 128**  
 139. **Figure 129**  
 140. **Figure 130**  
 141. **Figure 131**  
 142. **Figure 132**  
 143. **Figure 133**  
 144. **Figure 134**  
 145. **Figure 135**  
 146. **Figure 136**  
 147. **Figure 137**  
 148. **Figure 138**  
 149. **Figure 139**  
 150. **Figure 140**  
 151. **Figure 141**  
 152. **Figure 142**  
 153. **Figure 143**  
 154. **Figure 144**  
 155. **Figure 145**  
 156. **Figure 146**  
 157. **Figure 147**  
 158. **Figure 148**  
 159. **Figure 149**  
 160. **Figure 150**  
 161. **Figure 151**  
 162. **Figure 152**  
 163. **Figure 153**  
 164. **Figure 154**  
 165. **Figure 155**  
 166. **Figure 156**  
 167. **Figure 157**  
 168. **Figure 158**  
 169. **Figure 159**  
 170. **Figure 160**  
 171. **Figure 161**  
 172. **Figure 162**  
 173. **Figure 163**  
 174. **Figure 164**  
 175. **Figure 165**  
 176. **Figure 166**  
 177. **Figure 167**  
 178. **Figure 168**  
 179. **Figure 169**  
 180. **Figure 170**  
 181. **Figure 171**  
 182. **Figure 172**  
 183. **Figure 173**  
 184. **Figure 174**  
 185. **Figure 175**  
 186. **Figure 176**  
 187. **Figure 177**  
 188. **Figure 178**  
 189. **Figure 179**  
 190. **Figure 180**  
 191. **Figure 181**  
 192. **Figure 182**  
 193. **Figure 183**  
 194. **Figure 184**  
 195. **Figure 185**  
 196. **Figure 186**  
 197. **Figure 187**  
 198. **Figure 188**  
 199. **Figure 189**  
 200. **Figure 190**  
 201. **Figure 191**  
 202. **Figure 192**  
 203. **Figure 193**  
 204. **Figure 194**  
 205. **Figure 195**  
 206. **Figure 196**  
 207. **Figure 197**  
 208. **Figure 198**  
 209. **Figure 199**  
 210. **Figure 200**  
 211. **Figure 201**  
 212. **Figure 202**  
 213. **Figure 203**  
 214. **Figure 204**  
 215. **Figure 205**  
 216. **Figure 206**  
 217. **Figure 207**  
 218

TATGGGACCTTCAAC

TAHGGACCTCAGAC

C I C A I C C C C C A G A I  
 I I I I I I I I I I I I I I

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12

\_\_\_\_\_

[illegible]

```
QY 2670 AACCTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCAGCTGGCTGCAGCAGTG 2729
Db 1552 AACCTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCAGCTGGCTGCAGCAGTG 1611
QY 2730 GCTGAGCTGGCGCAGCAGCAGCTGGCTGCTC-TTCACAGTG-TCCGAGGCTGAGTGTGAG 2787
Db 1612 GCTGGAGCTGGCGCAGCAGCAGCTGGCTGCTC-TTCACAGTG-TCCGAGGCTGAGTGTGAG 1671
QY 2788 GCCGGCCAGGCGCCACACCTACACTCTCACCAGCTTTGGCA--CCCACCAAGGACAGGCA 2845
Db 1672 GCCGGCCAGG--CGAACACTACAAATTTACAGTTTGGGAACCCACCAAGGACAGGCA 1729
QY 2846 GAAGCCGGACAGGGGCGCTTCCCAACCGGGGAGA 2881
Db 1730 GAAGCCGGACAGGGGCTTTTCCCAACCGGGGAGA 1765

RESULT 12
AX054976
LOCUS AX054976 9700 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 91 from Patent WO0073328.
ACCESSION AX054976
VERSION AX054976.1 GI:12228344
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1. van Criekeing, W., Roelens, I., Bogaert, T. and Verwaerde, P.
   Unc-5 constructs and screening methods
   Patent: WO 0073328-A 91 07-DEC-2000;
   Devgen NV (BE)
FEATURES
    source
    1..9700
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /notes="plasmid pGC1037"
ORIGIN
    Query Match 44.8%; Score 1291.2; DB 6; Length 9700;
    Best Local Similarity 99.5%; Pred. No. 9e-204;
    Matches 1306; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1478 CACACTGCACCAAGCTCTCCACCTCTGAGCGCGAGGAGTTCGTCCCGCTCTCCAC 1537
Db 7 CACACTGCACCAAGCTCTCCACCTCTGAGCGCGAGGAGTTCGTCTCCCGCTCTCCAC 66

QY 1538 CCAGAACTACTTCGGCTCCCTGCGCCGAGGACCAACATGACCTATGGGACCTTCAA 1597
Db 67 CCAGAACTACTTCGGCTCCCTGCGCCGAGGACCAACATGACCTATGGGACCTTCAA 126

QY 1598 CTTCTCCGGGGCGGCTGATGATCCCTAATACAGTATCAGCTCTCCTATCCCCCAGAG 1657
Db 127 CTTCTCCGGGGCGGCTGATGATCCCTAATACAGTATCAGCTCTCCTATCCCCCAGAG 186

QY 1658 TGCCATACCCCGGAGGAAGATCTATGAGATCTACCTCAGCTGCAACGCGGAGACGT 1717
Db 187 TGCCATACCCCGGAGGAAGATCTATGAGATCTACCTCAGCTGCAACGCGGAGACGT 246

QY 1718 GAGTTTGCCCTTAGTGGCTGTGACACCTGTGATGCCATCGTTAGCTGTGGACCCCC 1777
Db 247 GAGTTTGCCCTTAGTGGCTGTGACACCTGTGATGCCATCGTTAGCTGTGGACCCCC 306

QY 1778 TGCGGCTGCTGCTACCCGCGCAGTATCTGCTATGACCACTGTGGGAGCCAGCC 1837
Db 307 TGCGGCTGCTGCTACCCGCGCAGTATCTGCTATGACCACTGTGGGAGCCAGCC 366

QY 1838 TGACAGCTGAGGCTGCGCTCAAAAAGCAGTCGTGCGAGGCGAGCTGGAGCAGATGT 1897
Db 367 TGACAGCTGAGGCTGCGCTCAAAAAGCAGTCGTGCGAGGCGAGCTGGG---AGGATGT 423

QY 1898 GTCGACCTGGCGGAGGAGCGCCCTCCACCTCTACTACTGCCAGCTGGAGCCAGTGC 1957
```

```
Db 424 GCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCAGCTGGAGGCGCAGTGC 483
QY 1958 CTGCTAGCTCTTCCACCGAGCAGCTGGGCGCGCTTTGGCCCTGTGGAGAGGCGCTCAGCGT 2017
Db 484 CTGCTAGCTCTTCCACCGAGCAGCTGGGCGCGCTTTGGCCCTGTGGAGAGGCGCTCAGCGT 543
QY 2018 GGTTCGCGCCAAAGCGCTCAAGCTGTCTTGTGTTTGGCGCGGTGCTGCACTCCCTCGA 2077
Db 544 GGTTCGCGCCAAAGCGCTCAAGCTGTCTTGTGTTTGGCGCGGTGCTGCACTCCCTCGA 603
QY 2078 GTACCAACATCGGGTCTACTGCTGCTGATGACACCCAGATGCACTCAAGAGGTGTGCA 2137
Db 604 GTACCAACATCGGGTCTACTGCTGCTGATGACACCCAGATGCACTCAAGAGGTGTGCA 663
QY 2138 GCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGA 2197
Db 664 GCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGA 723
QY 2198 CAGTTACCAACCTGGCGCTTATCCATCCAGATGTCAGCTCCCTGTGGAGAGTAA 2257
Db 724 CAGTTACCAACCTGGCGCTTATCCATCCAGATGTCGCCAGCTCCCTGTGGAGAGTAA 783
QY 2258 GCTCCTTGTGCTAGCTACCGAGGATCCCTTTTATCATCTGGAATGGCAGCGGTA 2317
Db 784 GCTCCTTGTGCTAGCTACCGAGGATCCCTTTTATCATCTGGAATGGCAGCGGTA 843
QY 2318 CTTGCACTGCACCTTCAACCTGGAGGCTGTCAGCCCGCAGCACTAGTACCTGCCCTGCAA 2377
Db 844 CTTGCACTGCACCTTCAACCTGGAGGCTGTCAGCCCGCAGCACTAGTACCTGCCCTGCAA 903
QY 2378 GCTGTGGTGTGCAAGTGGAGGCGCGAGGCTTTCAGCATCAACTTCAACATCAC 2437
Db 904 GCTGTGGTGTGCAAGTGGAGGCGCGAGGCTTTCAGCATCAACTTCAACATCAC 963
QY 2438 CAAGGACACAAAGTGTGCTGAGCTGCTGCTGAGAGTGAAGCGGGGCTCCAGCCCT 2497
Db 964 CAAGGACACAAAGTGTGCTGAGCTGCTGCTGAGAGTGAAGCGGGGCTCCAGCCCT 1023
QY 2498 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCCCTCACTTCGGCAGAAATATTTCCAGCCT 2557
Db 1024 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCCCTCACTTCGGCAGAAATATTTCCAGCCT 1083
QY 2558 GACCCACCTGTAGCGGGGTGCGAGCTGGCGGACTCTGGCCCGAGAACTCCACCTGGA 2617
Db 1084 GACCCACCTGTAGCGGGGTGCGAGCTGGCGGACTCTGGCCCGAGAACTCCACCTGGA 1143
QY 2618 CAGCCATCTCAGCTTCTTGGCTCCCAAGCCGAGCCCGCAGCATGATCTCAACTGTG 2677
Db 1144 CAGCCATCTCAGCTTCTTGGCTCCCAAGCCGAGCCCGCAGCATGATCTCAACTGTG 1203
QY 2678 GGAGGCGCGGCACTTCCCAAGCGGCAACCTCAGCCAGCTGGCTGAGAGTGGCTGACT 2737
Db 1204 GGAGGCGCGGCACTTCCCAAGCGGCAACCTCAGCCAGCTGGCTGAGAGTGGCTGACT 1263
QY 2738 GGGCCAGCAGACGCTGGCTCTTCAAGTGTGAGGCTGAGTGTGCTGAGG 2789
Db 1264 GGGCCAGCAGACGCTGGCTCTTCAAGTGTGAGGCTGAGTGTGCTGAGG 1315
```

```
RESULT 13
MMU72634
LOCUS MMU72634 9299 bp mRNA linear ROD 13-MAY-1997
DEFINITION Mus musculus rostral cerebellar malformation protein (rcm) mRNA,
complete cds.
ACCESSION U72634
VERSION U72634.1 GI:2088526
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 9299)
```

AUTHORS Ackerman,S.L., Kozak,L.P., Przyborski,S.A., Rund,L.A., Boyer,B.B. and Knowles,B.B.  
 TITLE The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein  
 JOURNAL Nature 386 (6627), 838-842 (1997)  
 MEDLINE 97271898  
 PUBMED 9126743  
 REFERENCE 2 (bases 1 to 9299)  
 AUTHORS Ackerman,S.L., Kozak,L.P., Rund,L.A. and Knowles,B.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA  
 FEATURES Location/Qualifiers  
 source  
 1..9299  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57B6/sJL"  
 /db\_xref="taxon:10090"  
 /chromosome="3"  
 /map="between Amy1 and Mpmv9"  
 1..9299  
 /gene="rcm"  
 151..2946  
 /gene="rcm"  
 /function="transmembrane receptor"  
 /note="similar to C. elegans UNC-5"  
 /codon\_start=1  
 /product="rostral cerebellar malformation protein"  
 /protein\_id="AA854103.1"  
 /db\_xref="GI:2088527"  
 /translation="MPKGLPAAACGLGIGYLLQMLVLPALALLSASGTGSAOODE  
 FFHELPFDSPEPLPHLEIPBEAYIVKNPVLNLYCKASPAIQIYFKCNSEVHQ  
 KDHVBERVDSGLIVREVSIEISQOQVEELFGPEDIWCQCVAMSGTTRKRYAV  
 RIAYLRTPFQELGKVELEQVLLQCRPEGIPVAEVLNKNEDIIIDPAEDRNFI  
 TIDHNLIIKQALSDTANYCTVAKNIIVAREKSTTIVIVYNGSWTTEWSVNCNR  
 GRGVQKTRCTNPAPLNGAFCEGOSQVQIACTTLCPCVDGRWTSNWSKSTCCTECTH  
 WRRECTAPAKPGKDKDGLVLQSKNCTDELQWAPDSDDVALVGIIVIAVTVCLA  
 ITVVVALFVTRKHDRFESDIIDSSALNGFQFQPNIKAAQDDLLAVPDDLTSAAMVR  
 GPVVALHDVSKIPMTNSPIIDELPNLKIIVYNSSGAVTFQDDDLAEFSSKLSQMTQS  
 LLEENALNKQSLARLTDSCFAFGFNLSLGGHLIIPNSGVSLILIPAGALPOGRVYE  
 MYVTVHRKNRPMEDSOTLLTPVSCGPPGALLTRPVILTLHHCADPSTEDWKTLQ  
 KNOAQGVQEDVVVGVNENFTPCYIQLDAEACHILNLSLTVLHVGQSTTKAAAKRL  
 KLAIFGLPCSSLSLEYSIRVYCLDDTDALKEVLQLEFRMGGLLEBPKALHFKFSLHN  
 LRUSIHLIAHSLKLLAKYQIPIPFYHWSGQRNLHCTFTLERLSLNTVELVKLC  
 VROVGEQGIQFLNCTVSEPTGIDPLDPASTITITVTPGSAPFSIPPIRQKLCSSL  
 DAPQTRGHDRMLAHKLNLDRYLNIFYATKSSPTGVLIDLWEAQNFPDGNLSMLAAVLE  
 EMGRHETVWSLAABGQY"

## ORIGIN

Query Match 32.6%; Score 940.2; DB 10; Length 9299;  
 Best Local Similarity 61.9%; Pred. No. 9.9e-146; Mismatches 56; Gaps 7;  
 Matches 1642; Conservative 0; Indels 943;

Qy 190 TGCCTGCTGCCAACCCGACCTGCTCCCACTTCCTGGTGGAGCCGCGAGGATGTGACA 249  
 Db 308 TTCCATCTGACCCACCTGAGCCATTGCCACATTCCTCATTTAGCCCCGAGGAGCTTACA 367  
 Qy 250 TCGTCAAGAAACAGCCAGTGTCTGTGTGTGTCAGAGCGCGTCCCGCCAGCAGATCTTCT 309  
 Db 368 TTGTGAAGAAACAGCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACT 427  
 Qy 310 TCAAGTCGACGGAGGTGGGTGGCCAGGTGGACCACTGATCGAGCGGACAGACG 369  
 Db 428 TCAAGTCGACAGCGAGGTGGGTTCATCAGAGGACCACTAGTAGACAGAGATGATG 487  
 Qy 370 GGAGCAGTGTGAGCCGACCATGGAGTCCCGATTATGTCTCAAGGCGAGGTGCGAGA 429  
 Db 488 AAACCTCTGGTCTAATTGTGAGAGAAGTGAGCATTGAGATTTTCAGCCGAGCGTGGAGG 547  
 Qy 430 AGGTGTTCCGCTGGAGGAATACTGGTCCAGTGGGTGGATGAGGTCTCTCGGGCAGCA 489  
 Db 548 AACTGTTGGTCCCTGAAGATTACTGGTCCAGTGTGGCTGGAGCTCAGCAGGACCTA 607

Qy 490 CCAAGAGTCAGAAAGGCTACATCCGCATAGCCAGATTGGCAAGAACTTCGAGCAGAGC 549  
 Db 608 CBAAGAGTCGAAGAGCATAGTTCGCAATTCGATTCGCGAAGACATTCGAGCAGAAC 667  
 Qy 550 CCGTGGCCAAAGAGGTGCTCCCTGGAGCAGGCAATCGTGTGCCCTGCCCTCCACCCGAGG 609  
 Db 668 CTTTGGGAAGAAAGTGTCTCTTGGAGCAGGAAGTCTTACTCCAGTGTCCGCAACCTGAAG 727  
 Qy 610 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGACCTGTGTGGACCCCTCCCTGG 669  
 Db 728 GGATCCCAAGTGGCTGAGGTGGAATGCTTAAAGNATGAAGACATAATGTATCTCTGCTGAAG 787  
 Qy 670 ACCCAATGTATACATCAACGCGGAGCAGACGCTGTGTGGCAGAGGCCCGCTTCTGTG 729  
 Db 788 ATCGGAATTTTATATTACTATCGATCAACAACCTGATCATCAAGCAAGCCGAGCTCTCAG 847  
 Qy 730 ACACGCCCAATACACTCGGTGGCCCAAGAACATCGTGGCAGCTCGCGCAGCGCTCCG 789  
 Db 848 ATACAGCAAAATATACCTGTGTGGCCAAAATATTGTTCACAGAGAAAAGACCAACAG 907  
 Qy 790 CTGCTGTCTACGTCTACGTGAACGGTGGGTGCTGACGTGGACCGGAGTGGTCCGTCTGCA 849  
 Db 908 CCACTGTCTACGTGTATGTTAATGTTGGTGTGCTGCTGACCTGGACAGAGTGTCTGTGTGTA 967  
 Qy 850 GCGCCAGCTGTGGGCGGGCTGGCAGAAACGGAGCCGAGCTGCACCAACCCGGGCGCTC 909  
 Db 968 ACAGCCGCTGTGGCGGAGGATATCAGAAACGCAAGAAACCTGCACCAACCCAGCCAC 1027  
 Qy 910 TCAACGGGGCGCTTCTGTGTGAGGGCAGATGTCATGACCCGACCGTCTCTCTCTGTC 969  
 Db 1028 TCAATGTTGGGGCTTCTGTGTGAGGGCAGAGTGTGCAGAA---AATAGCATGCACCTACGT 1084  
 Qy 970 TTGCTCTGTGGACGCGCAGCTGGAGCCGCTGGAGCAAGTGGTGGCGCTGTGGGCTGACT 1029  
 Db 1085 TATGTCAGTGGATGGTAGTGCACTTCATGGAGCAATGTTCACTGTGGGACTGAAT 1144  
 Qy 1030 GCACCCACTGGCGAGCGCTGAGTGTCTGACCCAGCACCCCCCAACGAGGGGAGGAGT 1089  
 Db 1145 GCACCCACTGGCGAGGAGGAGTGTACAGCACAGCCCCCAAGAACGGGGTAAAGACT 1204  
 Qy 1090 GCAGGGCACTGACTGGACACCGGCACTGTACCACTGACCTGTGTGTACACAGTGCCTT 1149  
 Db 1205 GTGATGGCTTGGCTCTCCAAATCCAAAGAACTGCATGATGGGCTGTGATGAGGCTGCTC 1264  
 Qy 1150 CTGSCCTGAGGAGTGGCCCTCTATCTGGG---CCTCATCGCGCTGGCGCTGTGCTCTGG 1206  
 Db 1265 CTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTATCGCTGTAACTGTCTGTCTGG 1324  
 Qy 1207 TCCTGTGCTGTGTCTCTCATCTCTGTTTATTGCGGAAGAGGAGGGGCTGGACTCAG 1266  
 Db 1325 CGATCACTGTTGTGGTGGCGCTGTTTGTGTATCGAAGAACACCGTGACTTTTGTGCTG 1384  
 Qy 1267 ATGTGGCTGACTCGTCCATTCTACCTCAGGCTTCAGCCCGCTCAGCATCAAGCCAGCA 1326  
 Db 1385 ACATCATTTGACTCCTCAGCACTCAATGGCGCTTTCAGCCTGTGAAATCAAGGCTGCCA 1444  
 Qy 1327 AAGCAGACAACCCCACTCTGCTCAACCATCCAGCCGACCTCAGCACCAACCAACCTACC 1386  
 Db 1445 GACAAATCTCTGGCTGTCCCCCTGACC-----TCACCTCAGCTGCAGCCATGTACA 1498  
 Qy 1387 AGGCACTCTCTGTCCCGCAGGATG-----GGCCAGCGCCCAAGTTCAGTCAACCAA 1441  
 Db 1499 GGGGACCTGTCTATGCTCTGCATGATGCTCAGACAAAATCCCAATACCAACTCTCCAA 1558  
 Qy 1442 TGGGCACTGCTAGCCCCCTGGTGGCGGCCACACACTGCACACACACTCTCCAC 1501  
 Db 1559 TTCTGGACCCCACTACCCCACTTGAAATTCAAAGTGTACAAAGCTCAGGTCTGTCACTC 1618  
 Qy 1502 CTC-----TGAGGCCGAGAGTTCGTCTCCCGCTCTCCACCC----- 1539  
 Db 1619 CTGAGGATGACCTTGGCGAGTTCCTCATCCAACTGTCCACCCAGATGACCCAGCTCTTGC 1678  
 Qy 1540 -----AGAACTACTTCCGCTCCCTCCCGCGGCGGACCCAGCA 1575

Db 1679 TAGAATAGAGCCCTTAACCTGAAGAACAGAGCTCGAAGACAGACTGACCATCCT 1738  
QY 1576 ACATGACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGTA 1635  
Db 1739 GCACAGATTGGTACCTTCAACTTCTTGGGGGTACCTCATCATTTCTAATTCAGGAG 1798  
QY 1636 TCAGCCTCTCATCCCCCAGATGCCATACCCAGGGAAGATCTATGAGATCTACCTCA 1695  
Db 1799 TAAGCTTGCTGATTCGCGCTGGGCCATTCCTCAGGGGAGAGTCTATGAATGTATGTA 1858  
QY 1696 CGCTGCACAGCCGAGAGAGTGAAGTTGCCCTAGCTGTGCTGTGACACCCCTGCTGATC 1755  
Db 1859 CTGTACACAGGAAGAATAATATGAGGCCGCCCATGGAAGACTCTCAGACCCCTACTTACCC 1918  
QY 1756 CCATCGTTAGCTGTGGACCCCTCGCTGCTGCTCACCGGCCAGTCTATCTGCTATGG 1815  
Db 1919 CTGTGCTGAGCTGTGGGCTCTCTGAGCTGTGCTGACCGGCCCTGCTATCTCATCTGC 1978  
QY 1816 ACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCCTCGGCTCAAAAAGCAGTGTGCG 1875  
Db 1979 ATCACTGTGAGACCCAGCAGCAGGACTGGAAGATCCAGCTCAAAAACAGGCACTGC 2038  
QY 1876 AGGCAGCTGGGAGCAGATGTGCTGCACTGGGCGAGAGGCGCCCTCCACCTCTACT 1935  
Db 2039 AGGACAAATGGGAGGATGTTGTGGTG---GTTGGGAGAGAACTTCACAAACCCCTGTT 2095  
QY 1936 ACTGCCAGCTGGAGGCGAGTGCCTGTCTACCTTTCACCGAGCAGCTGGGCGGCTTGGCC 1995  
Db 2096 ACATTAGCTGGATGAGAGGCTTGCCATATCTTCACAGAAACCTCAGTACCTATGCCC 2155  
QY 1996 TGTGGGAGGAGCCCTCAGCGTGGCTGCCGCCAGCGCTCAAGCTCTCTCTGTTGCGC 2055  
Db 2156 TGGTTGGGAGTCCACCAAGAGCAGTGCAGAGCTCTTAACTGGCCATCTTTGGGC 2215  
QY 2056 CGGTGCTGCACTGCTGCTGAGTACAAATCCGGGTCTACTGCTGATGACACCCACG 2115  
Db 2216 CCTCTGCTGCTCTTCCCTGGAGTACAGCATTAAGAGTCTACTGCTGGATGACACAGG 2275  
QY 2116 ATCACTCAAGAGGTGTGACCTGGAGAGCAGCTGGGGACAGCTGATCCAGAGC 2175  
Db 2276 ATGCTCTGAAGAGTCTTCAAACTGGAGGCAAAATGGGAGGACAGCTCTTGAAGAAG 2335  
QY 2176 CAGGCTCTGCACTCAAGGACAGTATACCAACCTGGCCCTATCCATCCAGATGTC 2235  
Db 2336 CCAAGGCTCTTCAATTTAAAGGAGCATCCAACTGGCCCTGTCTATTATGACATCG 2395  
QY 2236 CCAGCTCCCTGTGGAAGAGTAACTCTGTGAGTACAGGAGATCCCTTTTATCA 2295  
Db 2396 CCATTCCTCTGGAAGAGCAAAATGCTGGCTAAGTATCAGGAAATTCATTTTACCA 2455  
QY 2296 TCTGGAATGGCAGCAGCGGTACTTGCATGCACTTACCTTACCTGGAGCGTGTGAGCCCA 2355  
Db 2456 TCTGAGTGGCTCTCAAGAAACCTCACTGACCTTCACTCTGGAAGAGCTCAGCCTAA 2515  
QY 2356 GCATAGTACCTGGCTTCAAGCTGTGGGTGTGGAGGTGGAGGCGGAGCGGACAGCT 2415  
Db 2516 ACACAGTGAATCTGTTTCAAACTCTGTGTGGGAGGTTGAGAGAGAGGCGAGATCT 2575  
QY 2416 TCAGCATCACTTCAACATCACCAAGGACAAGGTTTCTGAGCTGTGGCTCTGGAGA 2475  
Db 2576 TCCAGCTCAACTGTACTGTGTGAGAGAACTTACTGGCATGCACTTACCTCTCTGGACC 2635  
QY 2476 GTGAAGCGGGGTCCAGCGCTGTGGGCCCGCAGTGCCTTCAAGATCCCTTCTCTATTC 2535  
Db 2636 CTGCTAGTACCATCACTGTCACCGGACCAAGTGTCTTTCAGCATTCCTCTCCCTATCC 2695  
QY 2536 GCGAGAAGATAATTTCCAGCTGGAGCCACCTGTAGGGGGGTGCGGAGTGGCGGACTC 2595  
Db 2696 GCGAGAAGTATGAGCAGCCTGGATGCCCTCAACAGAGGCGCATGACTGAGAGATGC 2755  
QY 2596 TGGCCCAAGAACTCCACCTGGACAGCATCTCAGCTTCTTTCCTTCCAGCCAGCCCA 2655

Db 2756 TGGCCCATAACTCAACTGGACAGGTACTTGAANTTACTTTGGCCACCAATTCGAGCCCAA 2815  
QY 2656 CAGCCATGATCTCAACTCTGTGGAGGCGGCGCACTTCCCCCAAGGCAACCTTCAGCCAGC 2715  
Db 2816 CTGCGCTAATCTCGATCTTTGGGAAGCACAGAACTTCCCAGATGGAACCTTCAGCATGC 2875  
QY 2716 TGSCTGACAGAGTGGCTGGACTGGCCAGCAGCAGCTGGCCTTTTCACAGTGTGGAGG 2775  
Db 2876 TGGCAGCGCTCTGGAAGAAATGGGAAGACATGACAGTGGTGTCTCTTGGCAGCAGAG 2935  
QY 2776 CTGAGTGTCTGA 2786  
Db 2936 GACAGTATTGA 2946

AV187310 2962 bp mRNA linear VRT 06-JUN-2003  
DEFINITION Gallus gallus UNC5-like protein 3 mRNA, complete cds.  
ACCESSION AY187310  
VERSION AY187310.1 GI:31442350  
KEYWORDS  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 2962)  
AUTHORS Guan,W. and Condic,M.L.  
TITLE Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during chick dorsal root ganglia development  
JOURNAL Gene Expr. Patterns 3, 369-373 (2003)  
REFERENCE 2 (bases 1 to 2962)  
AUTHORS Guan,W. and Condic,M.L.  
TITLE Direct Submission  
JOURNAL Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah, 20 North, 1900 East, Salt Lake City, UT 84132-3401, USA

FEATURES  
source  
1..2962  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9031"  
CDS  
1..2796  
/codon\_start=1  
/product="UNC5-like protein 3"  
/protein\_id="AA067275.1"  
/db\_xref="GI:31442351"  
/translation="MGKLGEGTAARCGLGMGVLLHSHVLPALAVLGASRPGSAQQDD  
FPEHLPETPPDPPEPLPHFLIEPREAYIVKNKPNLYCKASPATQIYFKCNSEBWHQ  
KDHVDERVDETSGLIVCEVSIIEISRQVVEELFGPEYWCQCVAMSSAGTTSRKAYV  
TIYLRKTFPEPLGKESLQEVLLQCRPEGIPIVAEVELKNEVIDPVEDRNFYI  
TIDHNLIIKARLSDTANYTCVAKNIIVAKRSTTATVIVVNGSTWEMWACNSRC  
GRGFORRTCTNPAPLNGAFCEGONVKIATCTLCPVDGKWTSMWSTCGTCTH  
WRRECTAPAKNGKDCGLVLOSNCCTDGLCMAAPSDVDVALIVGIVIAVCLA  
LSVVALFVYKXNRHDFPSDIIDSSALNGGFQPNVKAARDLLAVPDLTSAAMYR  
GPVALHDVSKIPMNSPILDPPLNLKIVNTSGAVTPODELSDFPSKLSQITQS  
LLENLTNVKQSLARQDTPSCFTAFNLSGLHLVINSVLSLIPAGAVPQGRVYE  
MYTVHRKEGRMPVEDSQTLLTPVSGPPGALLTRPWLTHHCAENMDMDWOIOL  
KHOAGQGWEDVVVVEENFTTPCYIOLDPEACHILLTETLSTVALVGSITKAAAKRL  
KLAIFGLSCSLEYSIRVCLDQDALKVELQLEROMGQGLLEPEKLPFGSTHN  
LRLSHDLAHLWLSKLPKAPQIEFFIHWGSCQRLHCTFTLFRFSLTLELVCKLC  
VRQVEGQIFOLNCSVSEBPTGIDYPIHMSAGSITTVGNFASIPFIPQKLGSL  
DAPQTRGHRMLAHLKLDRLYLFATKSPGTGVLDMWEAQNPDGNLSMLAAVLE  
EMGRHETVTVSAAAGNY"

ORIGIN  
Query Match 32.6%; Score 938.4; DB 5; Length 2962;  
Best Local Similarity 61.3%; Pred. No. 2.2e-145;  
Matches 1637; Conservative 0; Mismatches 961; Indels 72; Gaps 5;  
QY 190 TGCCTGGTGCACACCCGACCTGCTTCCCCACTTCTGCTGGAGCCGAGGATGTGTACA 249  
Db 158 TTCCTTGTGATCTCCAGAGCCATTCGCCCACTTCTCATTTGACCCCGAAGAGCTTACA 217







```
Db 2363 TCAATACCTGAGCTGCTGCAAACTCTGTGGGCAAGTCGAAGGAGGCGAGA 2422
Qy 2413 GCTTCAGCATCAACTTCAACATCACCAGGACACAAGGTTTGTGAGCTGTGCTCTGG 2472
Db 2423 TCTTCCAGCTGAATGCTCAGTATCAGAGGAAACCACTGGCAATTGATTATCCCATG 2482
Qy 2473 AGAGTGAAGCGGGGTCCCAAGCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCTCA 2532
Db 2483 ATTCAGCAGGAGCATCACACGATAGTTGGGCCCAACGCTTTCAGATCCCTTCCAA 2542
Qy 2533 TTGGGCGAAGATAATTTCCAGCTGTGACCCACCTGTAGGCGGGTGCAGACTGGGGA 2592
Db 2543 TAAGGCGAAGCTGTGAGCAGGCTGTGATGACCCAGACCCGCGGCGCATGACTGGAGGA 2602
Qy 2593 CTCTGGGCCAGAACTCCACTGAGACGACATCTCAGCTTCTTTTGGCTCCAGGCCAGCC 2652
Db 2603 TGCTGGGCCACAGCTGAATTTGACAGGTACCTAAATTTTGTACGAAAGTCGAGTC 2662
Qy 2653 CCACAGCATGATCTCAACTGTGGGAGCGCGCACTTCCCAACCGCAACCTCAGCC 2712
Db 2663 CCACTGGGGTGTCTGATCTCTGGGAGGCCAGATTTCCCTGATGGCACTGAGCA 2722
Qy 2713 AGCTGGCTGAGCAGTGGCTGGAGTGGCCAGCCAGCAGCTGGCTTTCACAGTGTGCG 2772
Db 2723 TGCTGGCAGCAGTGTCTGGAAGAAATGGAGCAGATGAACCGTTGTTCTTTGGCAGCAG 2782
Qy 2773 AGGCTGAGTGTGAGGCGCGCCAGGCCGA 2802
Db 2783 AAGGAATTAATGATGAGCCTTTGGCGTGA 2812
```

## RESULT 15

```
AB118026
LOCUS Rattus norvegicus Unc5h3 mRNA, complete cds.
DEFINITION Rattus norvegicus Unc5h3 mRNA, complete cds.
ACCESSION AB118026
VERSION AB118026.1 GI:40217509
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1
AUTHORS Kuramoto,T., Kuwamura,M. and Serikawa,T.
TITLE Rat neurological mutations cerebellar vermis defect and hobble are caused by mutations in the netrin-1 receptor gene Unc5h3
JOURNAL Mol. Brain Res. (2003) In press
REFERENCE 2 (bases 1 to 9328)
AUTHORS Kuramoto,T. and Serikawa,T.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-2003) Takashi Kuramoto, Institute of Laboratory Animals, Graduate School of Medicine, Kyoto University; Yoshidakocho-cho, Sakyo-ku, Kyoto 606-8501, Japan (E-mail:tkuramoto@anim.med.kyoto-u.ac.jp, URL:www.anim.med.kyoto-u.ac.jp, Tel:81-75-753-4494, Fax:81-75-753-4409)
```

## FEATURES

```
Location/Qualifiers
1..9328
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/chromosome="2"
gene 1..9328
/gene="Unc5h3"
CDS 151..2946
/gene="Unc5h3"
/codon_start=1
/protein_id="BAD05181.1"
/db_xref="GI:40217510"
/translation="MRKGLRATAARCGLSGLYLQMLVLPALALLSASGTSAOQDD
FFHELPEFPDPPEPLPHFLIEPEEAYIVKKNPVNLYKASPAQIYFKCNSEWVHQ
```

```
Query Match 32.5%; Score 936.8; DB 10; Length 9328;
Best Local Similarity 61.8%; Pred. No. 3.6e-145;
Matches 1631; Conservative 0; Mismatches 937; Indels 72; Gaps 6;

Qy 204 CCGGACCTGCTCCCACTTCTCCGAGCCGAGGATGTTACATCGTCAAGAACAAAG 263
Db 322 CCTGAGCCATTACACATTTCTCATTGAGCCCGAGGAGCTTACATTGAGAGAACAG 381

Qy 264 CAGTGTGCTTGTGTGAAGGCCGTGCCCGCCAGCAGATCTTCTTCAAGTGCAACGG 323
Db 382 CTTGTAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTTCTTCAAGTGCAACAGC 441

Qy 324 GAGTGGTGGCCAGTGGACACAGTGTGAGCCGAGCAGACAGCGGAGCAGTGTGTAG 383
Db 442 GAATGGTTCATCAGAAGGACCCAGTAGTAGAGAGGGGTAGATGAAACCTCCGTCTC 501

Qy 384 CCGACCATGAGGTCCGCAATTATGTCTCAAGCAGCAGTCCGAGAGGTGTTTCGGGCTG 443
Db 502 ATTGTGAGAGAGTGAAGATTTGAGATTTACGCCAGCAGTGTGAGGAGTGTTCGGGCTC 561

Qy 444 GAGGAATATGTTGTCAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 503
Db 562 GAAGATTACTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 621

Qy 504 GCCTACATCGCATAGCCAGATTGGCCAGAACTTCGAGCAGGAGCGCTGGCCAAAGGAG 563
Db 622 GCATAGTGGCATTCGCTATCTGCGGAAGACATTTGAGCAGGAGACCTTGGGAAAGAA 681

Qy 564 GTGTCTCTCGAGCAGGAGCATCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 623
Db 682 GTGTCTCTCGAGCAGGAGTGTATCTCAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 741

Qy 624 GAGTGTGAGTGTCTCGGAAACAGGAGCTGTGTGAGCCCGTCCCTCGAGCCCAATGTATAC 683
Db 742 GAGTGTGAGTGTCTAAAGAAATGAAGACATAATTGATCTCTGTCGAAGATTCGGAATTTAT 801

Qy 684 ATCAGCGGAGCAGCAGCTGTGTGTGTCGAGCGCCGCTGTCGTCGTCGTCGTCGTCGTCG 743
Db 802 ATCAGTATGATCAACCTGATCATCAAGCAGGAGCAGCTCTCAGATACAGCGAATTAT 861

Qy 744 ACCTGCTGGCCCAAGAAACATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 803
Db 862 ACCTGCTGTGCAAAAACATTTGTCGAGGAGAAAGACCAACAGCAGCCGCTCATCTGTG 921

Qy 804 TACGTGAACGTTGGTGTGTGTCGAGCGTGGACCGAGTGTGTGTCGTCGTCGTCGTCGTCG 863
Db 922 TATGTTAATGTTGGTGTGTGTCGAGCTGGCAGNATGTTGTGTGTGTAAAGCGCTGTGGG 981

Qy 864 CCGGCTGGCAGAAACGAGCGGAGCTGTGACCAACCCGCGGCGCTCTCTCAACGCGGGCGCT 923
```

/gene="Unc5h3"

/note="simple sequence repeats"

/rpt\_unit="gt"

9309..9314

/gene="Unc5h3"

## ORIGIN

Query Match

Best Local Similarity

Matches 1631;

Conservative

0; Mismatches

937; Indels

72; Gaps

6;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 09:36:51 ; Search time 1081 seconds  
(without alignments)  
11321.978 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccggctgag.....ccttccccacacgggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2881	100.0	2881	6	Abk49422 DNA encod
2	2676.4	92.9	2752	6	Abk37922 CDNA enco
3	2615	90.8	2697	6	Abk52891 Human net
4	2402.6	83.4	3580	6	Abk15169 Human REP
5	2296.2	79.7	2907	4	Aak52261 Human pol
6	2235.8	77.6	3014	2	Aav52940 Rat UNC-5
7	2189.4	76.0	2697	6	Aas16843 Rat netri
8	1570.4	54.5	1787	2	Aav52941 Human UNC
9	1291.2	44.8	9700	4	Aac90958 Plasmid p
10	1193.6	41.4	1321	4	Aah99530 Human pro
11	916	31.8	3646	5	Aas75738 DNA encod
12	889	30.9	2860	6	Abt06279 Human NOV
13	887.4	30.8	2860	6	Abt06280 Human NOV
14	875.8	30.4	2995	6	Abk92062 DNA encod
15	866.4	30.1	2895	6	Abg93898 Human tra
16	853.6	29.6	3884	4	Aas21316 Human cDN
17	853.6	29.6	3884	7	ACd23925 Novel hum
18	853.6	29.6	3884	7	ACA67066 CDNA enco
19	853.6	29.6	3884	7	ACA30375 CDNA enco
20	853.6	29.6	3884	7	Abx89213 DNA encod
21	853.6	29.6	3884	7	ACd41867 Human sec
22	853.6	29.6	3884	7	ACA04096 Human cDN
23	853.6	29.6	3884	8	Ada45664 Novel hum

24	853.6	29.6	3884	8	ADA76095 Human PRO
25	853.6	29.6	3884	8	ADA18745 Human PRO
26	853.6	29.6	3884	8	ADA61368 Homo sapi
27	853.6	29.6	3884	8	ADB19153 Novel hum
28	853.6	29.6	3884	8	ADB27694 CDNA enco
29	853.6	29.6	3884	8	ADA86173 Novel hum
30	853.6	29.6	3884	8	ADB15737 Human PRO
31	853.6	29.6	3884	8	ADA47523 Human PRO
32	853.6	29.6	3884	8	ADA67318 Human PRO
33	853.6	29.6	3884	8	ADB30325 CDNA enco
34	853.6	29.6	3884	8	ADA85621 Novel hum
35	853.6	29.6	3884	8	ADA96833 Human PRO
36	853.6	29.6	3884	8	ADA79137 Human PRO
37	853.6	29.6	3884	8	ADA87276 Novel hum
38	853.6	29.6	3884	8	ADB16478 Human PRO
39	853.6	29.6	3884	8	ADA91570 Novel hum
40	853.6	29.6	3884	8	ADB14633 Human PRO
41	853.6	29.6	3884	8	ADB14633 Human PRO
42	853.6	29.6	3884	8	ADB18594 Novel hum
43	853.6	29.6	3884	8	ADA93809 Human PRO
44	853.6	29.6	3884	8	ADB19705 Novel hum
45	853.6	29.6	3884	8	ADB13017 Human PRO
					ACd98496 Novel hum

## ALIGNMENTS

### RESULT 1

ABK49422  
ID ABK49422 standard; DNA; 2881 BP.  
XX AC ABK49422;  
XX AC ABK49422;  
DT 15-JUL-2002 (first entry)  
XX DE DNA encoding human UNC5-like protein NOV1.  
XX DE Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;  
XX KW cell signal processing; metabolic pathway modulation; cancerous tissue;  
XX KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;  
XX KW chromosome 13; gene; ds.  
XX OS Homo sapiens.  
FH Key Location/Qualifiers  
CDS 87..2786  
FT /\*tag= a  
FT /product= "Human UNC5-like protein NOV1"  
XX

WG200229038-A2.

11-APR-2002.

04-OCT-2001; 2001WO-US031377.

04-OCT-2000; 2000US-0237862P.

(CURA-) CURAGEN CORP.

Herrmann JL, Rastelli L, Shinkets RA;

WPI; 2002-340104/37.

P-PSDB; AAU79939.

Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for treating cardiomyopathy, atherosclerosis, and cancer.

Claim 8; Page 7-8; 180pp; English.

The present invention relates to a new NOVX polypeptide having a 900 (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6) residue amino acid sequence, as given in the specification. The novel polypeptide, and its encoding polynucleotide, are used to treat

	CC	cardiomyopathy, atherosclerosis, cancer or a disease related to cell
CC	CC	signal processing and metabolic pathway modulation, in a human. Detecting
CC	CC	the polypeptide or polynucleotide is useful for identifying cancerous
CC	CC	tissue. The antibody can be used to treat diabetes or cancer. The host
CC	CC	cells can be used to produce non-human transgenic animals useful in drug
CC	CC	screening. The present nucleic acid sequence is that of the human UNC5-
CC	CC	like NOV1 gene located on chromosome 13. This sequence encodes the human
CC	CC	UNC5-like protein NOV1 of the invention
XX		
QQ	Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 2881; DB 6; Length 2881;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGCCCGCGCCCG 60
DB	1	AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGCCCGCGCCCG 60
QY	61	GCCCGCCGCTGCCCGCCCGCGGCCATGGCGTCCGCGCCGCGCCCTGTGGCCAGCGTCC 120
DB	61	GCCCGCCGCTGCCCGCCCGCGGCCATGGCGTCCGCGCCGCGCCCTGTGGCCAGCGTCC 120
QY	121	TGGCATAGTCTCTCGCCGCTTGGCTCCGCGGCTCGGGTGCGCCAGCAGATGCCACCGTGG 180
DB	121	TGGCATAGTCTCTCGCCGCTTGGCTCCGCGGCTCGGGTGCGCCAGCAGATGCCACCGTGG 180
QY	181	CCAAACCCAGTGCCTGGTGCCAAACCCGAGACTGTCTCCCACTTCTCTGGTGGAGCCGAGG 240
DB	181	CCAAACCCAGTGCCTGGTGCCAAACCCGAGACTGTCTCCCACTTCTCTGGTGGAGCCGAGG 240
QY	241	ATGTGTACATCGTCAAGAACAGCCAGTGTCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
DB	241	ATGTGTACATCGTCAAGAACAGCCAGTGTCTTGTGTGCAAGGCCGTGCCCGCCACGC 300
QY	301	AGATCTTCTTCAAGTGCAAACGGGGAGTGGGTGCGCCAGTGGACCACTGTATCGAGCGCA 360
DB	301	AGATCTTCTTCAAGTGCAAACGGGGAGTGGGTGCGCCAGTGGACCACTGTATCGAGCGCA 360
QY	361	GCACAGACGGAGCAGTGGTGAGCCGACATGGAGGTCCGCACTAACTCTCAAGGCAGC 420
DB	361	GCACAGACGGAGCAGTGGTGAGCCGACATGGAGGTCCGCACTAACTCTCAAGGCAGC 420
QY	421	AGGTGCGAAGGTGTTCCGGCTGGAGGAATACTGGTGCCAGTGGCTGGCATGGAGCTCCT 480
DB	421	AGGTGCGAAGGTGTTCCGGCTGGAGGAATACTGGTGCCAGTGGCTGGCATGGAGCTCCT 480
QY	481	CGGGCACCAACGAAGTGCAGAGGCCATACATCCGCATAGCCAGATGGCGCAGACTTCG 540
DB	481	CGGGCACCAACGAAGTGCAGAGGCCATACATCCGCATAGCCAGATGGCGCAGACTTCG 540
QY	541	AGCAGGACCGCTGCCCAAGGAGGTGTCCTGGAGCAGGGCATCGTGTGCCCTGCCGTC 600
DB	541	AGCAGGACCGCTGCCCAAGGAGGTGTCCTGGAGCAGGGCATCGTGTGCCCTGCCGTC 600
QY	601	CACCGGAGGGCATCCTCCAGCCGAGGTGGAGTGGCTCCGGAACAGAGCACTGTGTGGACC 660
DB	601	CACCGGAGGGCATCCTCCAGCCGAGGTGGAGTGGCTCCGGAACAGAGCACTGTGTGGACC 660
QY	661	CGTCCCTGGACCCCAATGTATACATCAGCGGGACACAGCTGGTGGTGCGACAGGCC 720
DB	661	CGTCCCTGGACCCCAATGTATACATCAGCGGGACACAGCTGGTGGTGCGACAGGCC 720
QY	721	GCCCTTGTGACACGGCCAACTACACTCGCTGGCGCCAAAGAACATCGTGGCACATCGCCGCA 780
DB	721	GCCCTTGTGACACGGCCAACTACACTCGCTGGCGCCAAAGAACATCGTGGCACATCGCCGCA 780
QY	781	GGCGCTCCGCTGTGTTCATCGTCTACGTGAAACGGTGGGTGGTTCAGCTGGACCGAGTGGT 840
DB	781	GGCGCTCCGCTGTGTTCATCGTCTACGTGAAACGGTGGGTGGTTCAGCTGGACCGAGTGGT 840
QY	841	CGTGTTCGACGGCCAGCTGTGGCGCGGGTGGCAGAAACGGAGCCGAGACTGCACCAACC 900

```
QY 1981 TGGGCGCTTTTGGCTTGGTGGAGAGCCCTCAGCGTGGCTGCCGCAAGCGCTCAAGC 2040
Db 1981 TGGGCGCTTTTGGCTTGGTGGAGAGCCCTCAGCGTGGCTGCCGCAAGCGCTCAAGC 2040
QY 2041 TGGCTTCTGTTGGCGCGGTGGCTTGGAGTCAAACTCCCTCGAGTCAAACTCCGGTCTACTGCC 2100
Db 2041 TGGCTTCTGTTGGCGCGGTGGCTTGGAGTCAAACTCCCTCGAGTCAAACTCCGGTCTACTGCC 2100
QY 2101 TGCATGACACCCAGCATGCACTCAAGAGAGTGGTCAAGTGGAGAACAGTGGGGGAC 2160
Db 2101 TGCATGACACCCAGCATGCACTCAAGAGAGTGGTCAAGTGGAGAACAGTGGGGGAC 2160
QY 2161 AGCTGATCCAGGAGCCACCGGTCTGTCACATTCAGGAGACAGTTACCAACAACCTGGCGCTAT 2220
Db 2161 AGCTGATCCAGGAGCCACCGGTCTGTCACATTCAGGAGACAGTTACCAACAACCTGGCGCTAT 2220
QY 2221 CAGTCCAGATGTGCCAGTCCCTGTGGAAGAGTAAAGCTTCCTGTGAGTACCAAGGAGA 2280
Db 2221 CAGTCCAGATGTGCCAGTCCCTGTGGAAGAGTAAAGCTTCCTGTGAGTACCAAGGAGA 2280
QY 2281 TCCCTTTTATCATCTGGAATGGACGCGGTACTTTGCACTGCACTTCCCTGTGAGTACCAAGGAGA 2340
Db 2281 TCCCTTTTATCATCTGGAATGGACGCGGTACTTTGCACTGCACTTCCCTGTGAGTACCAAGGAGA 2340
QY 2341 AGCGTGTCAAGCCAGCACTAGTGAAGTGGCTGCAAGCTGTGGTGTGGAGTGGAGG 2400
Db 2341 AGCGTGTCAAGCCAGCACTAGTGAAGTGGCTGCAAGCTGTGGTGTGGAGTGGAGG 2400
QY 2401 GCGACGGGAGAGCTTCAAGTCAAACTTCAATCAACATCAACAGGACACAAGTTTGTGTGAGC 2460
Db 2401 GCGACGGGAGAGCTTCAAGTCAAACTTCAATCAACATCAACAGGACACAAGTTTGTGTGAGC 2460
QY 2461 TGTGTGCTCTGGAGAGTGAAGCGGGGTCCAGCCTGTGGTGGCCCGGAGTGCCTTCAAGA 2520
Db 2461 TGTGTGCTCTGGAGAGTGAAGCGGGGTCCAGCCTGTGGTGGCCCGGAGTGCCTTCAAGA 2520
QY 2521 TCCCTTCTCTCAATTCGGGAGAGATATTTCCAGCCTGGACCCACCTGTAGGCGGGGTG 2580
Db 2521 TCCCTTCTCTCAATTCGGGAGAGATATTTCCAGCCTGGACCCACCTGTAGGCGGGGTG 2580
QY 2581 CGGACTGGCGGACTTGGGCCAGAAAATCACTGACAGCCTTCACTTCTTTTGCCT 2640
Db 2581 CGGACTGGCGGACTTGGGCCAGAAAATCACTGACAGCCTTCACTTCTTTTGCCT 2640
QY 2641 CCAAGCCAGCCAGCCAGCAGCAGTCACTCAAGTGTGGAGGGCGGACCTTCCCAAG 2700
Db 2641 CCAAGCCAGCCAGCCAGCAGCAGTCACTCAAGTGTGGAGGGCGGACCTTCCCAAG 2700
QY 2701 GCAACTCAGCCAGCTGGCTGAGCAGTGGCTGGAAGTGGCCAGCAGCAGCTGGGCTCT 2760
Db 2701 GCAACTCAGCCAGCTGGCTGAGCAGTGGCTGGAAGTGGCCAGCAGCAGCTGGGCTCT 2760
QY 2761 TCACAGTGTGAGGAGTGTGAGTGGCGGCGGCGGAGCCGACACCTACCTTCAACAG 2820
Db 2761 TCACAGTGTGAGGAGTGTGAGTGGCGGCGGCGGAGCCGACACCTACCTTCAACAG 2820
QY 2821 CTTTGGCACCACCAAGGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2880
Db 2821 CTTTGGCACCACCAAGGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2880
QY 2881 A 2881
Db 2881 A 2881
```

## RESULT 2

ABK37922

ID ABK37922 standard; cDNA; 2752 bp.

XX

AC ABK37922;

XX

DT 21-MAY-2002 (first entry)

XX cDNA encoding Human protein NOV1.

Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes; cell signal processing disorder; metabolic disorder; obesity; infection; anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension; osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris; myocardial infarction; ulcer; allergy; benign prostatic hypertrophy; psychosis; neurological disorder; anxiety; schizophrenia; manic depression; dementia; dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; gene therapy.

XX Homo sapiens.

XX WO200210216-A2.

XX 07-FEB-2002.

XX 30-JUL-2001; 2001WO-US024225.

XX 28-JUL-2000; 2000US-0221409P.

XX 04-AUG-2000; 2000US-0222840P.

XX 04-AUG-2000; 2000US-0223752P.

XX 04-AUG-2000; 2000US-0223762P.

XX 04-AUG-2000; 2000US-0223769P.

XX 04-AUG-2000; 2000US-0223770P.

XX 14-AUG-2000; 2000US-0225146P.

XX 15-AUG-2000; 2000US-0225392P.

XX 15-AUG-2000; 2000US-0225470P.

XX 01-FEB-2001; 2000US-0263662P.

XX 05-APR-2001; 2001US-0281645P.

XX (CURA-) CURAGEN CORP.

XX Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;

XX Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;

XX WPI; 2002-180074/23.

XX P-PSDB; AAU85403.

XX New isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide, useful for treating cardiomyopathy, atherosclerosis, infections, cancer, neurodegenerative, metabolic, hematopoietic and immune disorders.

XX Claim 9; Page 9-10; 213pp; English.

XX The invention relates to an isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature form. Also included are the nucleic acids encoding the NOVX proteins, a vector comprising the nucleic acid, a cell comprising the vector, an anti-NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the antibody are useful for treating or preventing a NOVX-associated disorder, where the disorder is selected from cardiomyopathy, atherosclerosis, diabetes, a disorder related to cell signal processing and metabolic pathway modulation, metabolic disorders, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, and the various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases, bacterial, fungal, protozoal and viral infections, pain, bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's disease, multiple sclerosis, Albright Hereditary Osteodysplasia, Pectoris, myocardial infarction, ulcer, allergy, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, and dyskinesias, such as Huntington's disease and Gilles de la Tourette's syndrome. The nucleic acid is useful in gene therapy. The present sequence encodes a NOVX protein











Db 2635 CGAGTGGCTGACTGGGCGCAGCAGCTGCTTTCACAGTGTGGAGCTGAGTGC 2694  
 Qy 2784 TGA 2786  
 Db 2695 TGA 2697  
 RESULT 4  
 ABK15169  
 ID ABK15169 standard; DNA; 3580 BP.  
 AC ABK15169;  
 XX 23-APR-2002 (first entry)  
 DT  
 XX Human REPTR 1 cDNA sequence.  
 DE  
 XX REPTR; human; antiinflammatory; cytostatic; immunosuppressive; antiviral;  
 KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;  
 KW antiallergic; antibody; immunogen; endometriosis;  
 KW gastrointestinal disorder; gastritis; oesophageal carcinoma;  
 KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;  
 KW endocrine disorder; hypothyroidism; rheumatoid arthritis; allergy;  
 KW autoimmune disease; inflammatory disease; infertility; receptor;  
 KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;  
 KW osteoarthritis; diabetes mellitus; multiple sclerosis;  
 KW systemic lupus erythematosus; cell proliferative disorder; cancer;  
 KW developmental disorder; Duchenne muscular dystrophy; gene;  
 KW Becker muscular dystrophy; neurological disorder; epilepsy;  
 KW Alzheimer's disease; Huntington's disease; reproductive disorder; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 4..2532  
 FT /\*tag= a  
 FT /product= "REPTR1 protein"  
 XX  
 XX WO200198354-A2.  
 XX  
 XX 27-DEC-2001.  
 XX  
 XX 21-JUN-2001; 2001WO-US019942.  
 XX  
 XX 21-JUN-2000; 2000US-0214027P.  
 PR 25-AUG-2000; 2000US-0228045P.  
 PR 12-DEC-2000; 2000US-0255104P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 XX Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT,  
 PI Lal P, Policky JL, Azinzai Y, Lu DM, Gaul R, Yao MG, Burford N,  
 PI Hafalia AJA, Baughn MR, Bandman O, Patterson C, Yang J, Xu Y,  
 PI Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;  
 XX  
 DR WPI; 2002-090432/12.  
 DR P-PSDB; AAU17818.  
 XX  
 XX Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in  
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
 PT gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell  
 PT proliferative (e.g. cancer) disorders.  
 XX  
 PS Claim 57; Page 142-143; 157pp; English.  
 XX  
 CC This invention relates to twelve human receptors cDNA sequences referred  
 CC to as REPTR-1 to REPTR-12), and the proteins encoded thereby. The  
 CC proteins of the invention may have antiinflammatory, cytostatic,  
 CC immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active  
 CC general, anticonvulsant, nootropic, neuroprotective, antiallergic  
 CC activities. The sequences of the invention may be used to produce REPTR  
 CC agonists or antagonists, and the protein sequences may be used to raise

CC anti-REPTR antibodies. These molecules and the REPTR polynucleotides and  
 CC polypeptides of the invention are useful in the diagnosis, treatment and  
 CC prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,  
 CC Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine  
 CC (e.g. hypothyroidism, Kallman's disease), autoimmune/inflammatory  
 CC (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,  
 CC allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,  
 CC systemic lupus erythematosus), cell proliferative (e.g. cancer),  
 CC developmental (e.g. Duchenne and Becker muscular dystrophy), neurological  
 CC (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and  
 CC reproductive (e.g. infertility, endometriosis) disorders. Numerous other  
 CC examples of each disorder are given in the specification. The present  
 CC sequence represents the human REPTR1 cDNA sequence of the invention  
 XX  
 SQ Sequence 3580 BP; 670 A; 1233 C; 1025 G; 652 T; 0 U; 0 Other;  
 Query Match 83.4%; Score 2402.6; DB 6; Length 3580;  
 Best Local Similarity 93.4%; Pred. No. 0;  
 Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;  
 Qy 84 GCATGCGGCTCGGGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCTTGG 143  
 Db 1 GCATGCGGCTCGGGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCTTGG 60  
 Qy 144 CTCGCGGCTCGGGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCTTGG 203  
 Db 61 CTCGCGGCTCGGGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCTTGG 120  
 Qy 204 CCGGACCTGCTCCCACTTCTCTGGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCG 263  
 Db 121 CCGGACCTGCTCCCACTTCTCTGGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCG 180  
 Qy 264 CAGTGTCTGCTTGTGTGCAAGCGCTGCGCGCCAGCAGATCTTCTCAAGTGCACCGG 323  
 Db 181 CAGTGTCTGCTTGTGTGCAAGCGCTGCGCGCCAGCAGATCTTCTCAAGTGCACCGG 240  
 Qy 324 GAGTGGTGGCCAGAGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCTTGG 383  
 Db 241 GAGTGGTGGCCAGAGTGGCCAGAGTGGCCAGCGCTCTGGGCGATAGTCTCGCGCTTGG 300  
 Qy 384 CCGACCATGAGGTGCGCATTAATGTCCTCAAGCAGCAGGTTCGAGAGTGTTCGGGCTG 443  
 Db 301 CCGACCATGAGGTGCGCATTAATGTCCTCAAGCAGCAGGTTCGAGAGTGTTCGGGCTG 360  
 Qy 444 GAGGAATCTGTTGTCAGTGGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 503  
 Db 361 GAGGAATCTGTTGTCAGTGGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATG 420  
 Qy 504 GCCTACATCCGATAGCCAGATTGCGCAAGACTTTCGAGCAGAGCGCTTGGCCAAAGGAG 563  
 Db 421 GCCTACATCCGATAGCCAGATTTTCGCAAGAACTTCGAGCAGAGCGCTTGGCCAAAGGAG 480  
 Qy 564 GTGTCTCTGAGCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623  
 Db 481 GTGTCTCTGAGCAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 Qy 624 GAGTGGAGTGGCTCGGAAACGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 683  
 Db 541 GAGTGGAGTGGCTCGGAAACGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
 Qy 684 ATCAGCGGGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 743  
 Db 601 ATCAGCGGGAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 Qy 744 ACCTGCTGGCCAAAGAAACATGCTGCGAGCTGCGCGCAGCGCTTCTGCTGCTGCTGCTGCTG 803  
 Db 661 ACCTGCTGGCCAAAGAAACATGCTGCGAGCTGCGCGCAGCGCTTCTGCTGCTGCTGCTGCTG 720  
 Qy 804 TACGTGAACGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 863  
 Db 721 TAC-----  
 Qy 864 CGCGCTGGCAGAAACGAGCGGAGTGTGCACCAACCGCGCGCTTCTCAACGGGGCGCT 923

D	b	724	-----	723	D	b	1687	CAGCTGGAGGCGCCTGCTACGTCTCCACGAGACGCTGGGCGCTTTGCCCTGGTG	1746
Q	y	924	TTCTGTGAGGGGAGAAATGTCCATGACGACCGTCTCTCTGCTGTGTCTGTGGAC	983	Q	y	2001	GGAGAGGCGCTCAGCGTGGCTCGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTG	2060
D	b	724	-----	729	D	b	1747	GGAGAGGCGCTCAGCGTGGCTCGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTG	1806
Q	y	984	GGCAGCTGGAGCCCGTGGAGCAAGTGTGGGCTGTGGGCTGACCTGACCCACTGGCGG	1043	Q	y	2061	GCTGTGACCTCCTCGAGTACAACTCCGGGTCTACTGCTGTCATGACACCCACGATGCA	2120
D	b	730	GGCAGCTGGAGCCCGTGGAGCAAGTGTGGGCTGTGGGCTGACCTGACCCACTGGCGG	789	D	b	1807	GCTGTGACCTCCTCGAGTACAACTCCGGGTCTACTGCTGTCATGACACCCACGATGCA	1866
Q	y	1044	AGCGGTGAGTCTGTACCCAGCACCCCGCAACGGAGGGGAGAGTGGCAGGGCACTGAC	1103	Q	y	2121	CTCAAGAGAGTGTGAGTGGAGAGAGTGGGGGGGACAGCTGATCAGAGGACCCACGG	2180
D	b	790	AGCGGTGAGTCTGTACCCAGCACCCCGCAACGGAGGGGAGAGTGGCAGGGCACTGAC	849	D	b	1867	CTCAAGAGAGTGTGAGTGGAGAGAGTGGGGGGGACAGCTGATCAGAGGACCCACGG	1926
Q	y	1104	CTGGACACCGCAACTGTACAGTGTACCTGTGTACACAGTGTCTTGTGSCCTTGAGGAC	1163	Q	y	2181	GTCTGTGACTTCAAGAGACAGTTACCAACCTCGCTTATCCATCCACGATGTGCCGAGC	2240
D	b	850	CTGGACACCGCAACTGTACAGTGTACCTGTGTACACAGTGTCTTGTGSCCTTGAGGAC	909	D	b	1927	GTCTGTGACTTCAAGAGACAGTTACCAACCTCGCTTATCCATCCACGATGTGCCGAGC	1986
Q	y	1164	GTGGCCCTCTATGTGGGCTCATGCGCGTGGCGGTGTGCTGTGCTGTGCTGTGCTGTC	1223	Q	y	2241	TCCCTGTGGAAGAGTAAAGCTCCTTGTGTCAGTACACAGAGATCCCTTTTATCACTGG	2300
D	b	910	GTGGCCCTCTATGTGGGCTCATGCGCGTGGCGGTGTGCTGTGCTGTGCTGTGCTGTC	969	D	b	1987	TCCCTGTGGAAGAGTAAAGCTCCTTGTGTCAGTACACAGAGATCCCTTTTATCACTGG	2046
Q	y	1224	CTCATCTCTCTTATTCGCGAAGAGAGGGGCTGACTCAGATGTGGTGTGCTGTGTC	1283	Q	y	2301	AATGGACGCGAGCGGTACTTGTGACTGACCTTCAACCTGGAGCGGTGTCAGCCCGACACT	2360
D	b	970	CTCATCTCTCTTATTCGCGAAGAGAGGGGCTGACTCAGATGTGGTGTGCTGTGTC	1029	D	b	2047	AATGGACGCGAGCGGTACTTGTGACTGACCTTCAACCTGGAGCGGTGTCAGCCCGACACT	2106
Q	y	1284	ATTCTACCTCAGGCTTCCAGCCGCTCAGCATCAAGCCAGCAAGAGAGAGAGAGAGAG	1343	Q	y	2361	AGTGAAGTGGCTGCAAGCTGTGGGTGTGGCAGTGGAGGGCGACGGGCGAGAGCTTCAGC	2420
D	b	1030	ATTCTACCTCAGGCTTCCAGCCGCTCAGCATCAAGCCAGCAAGAGAGAGAGAGAG	1089	D	b	2107	AGTGAAGTGGCTGCAAGCTGTGGGTGTGGCAGTGGAGGGCGACGGGCGAGAGCTTCAGC	2166
Q	y	1344	CTGCTACCAATCCAGCCGAGACTCAG---CAGCAACCACTTACAGGGGAGTCTCTGT	1400	Q	y	2421	ATCAACTTCAACATCAACCAAGGACACAAAGTGTGCTGAGCTGTGGGTCTGTGGAGGTGAA	2480
D	b	1090	CTGCTACCAATCCAGCCGAGACTCAGCAACCACTTACAGGGGAGTCTCTGT	1149	D	b	2167	ATCAACTTCAACATCAACCAAGGACACAAAGTGTGCTGAGCTGTGGGTCTGTGGAGGTGAA	2226
Q	y	1401	CCCCGGCAGAGTGGGCGGAGCCCAAGTTCAGCTCACCAGTGGGCACTGTCTCAGGCCC	1460	Q	y	2481	GGGGGGTCCAGCCCTGTGGTGGGCGGCGCTTCAAGATCCCTTCCCTCATTTCCGCGAG	2540
D	b	1150	CCCCGGCAGAGTGGGCGGAGCCCAAGTTCAGCTCACCAGTGGGCACTGTCTCAGGCCC	1209	D	b	2227	GGGGGGTCCAGCCCTGTGGTGGGCGGCGCTTCAAGATCCCTTCCCTCATTTCCGCGAG	2286
Q	y	1461	CTGGGTGGGCGGCGGACACACTGACACAGCTCTCCACCTCTGAGGGCGAGGATTC	1520	Q	y	2541	AGGATATTTCCAGCTGGACCCACACCTGTAGGGGGGTGCCGACTGGCGGACTCTGGCC	2600
D	b	1210	CTGGGTGGGCGGCGGACACACTGACACAGCTCTCCACCTCTGAGGGCGAGGATTC	1269	D	b	2287	AGGATATTTCCAGCTGGACCCACACCTGTAGGGGGGTGCCGACTGGCGGACTCTGGCC	2346
Q	y	1521	GTCTCCGCGCTCTCCACCGAGAACTATCTTCGCTCCCTGCGGCGAGGACAGCAATG	1580	Q	y	2601	CAGAAAATCCACCTGGACAGCCATCTCAGCTCTTTTCCCTCCAAAGCCGAGCCCGACAGCC	2660
D	b	1270	GTCTCCGCGCTCTCCACCGAGAACTATCTTCGCTCCCTGCGGCGAGGACAGCAATG	1329	D	b	2347	CAGAAAATCCACCTGGACAGCCATCTCAGCTCTTTTCCCTCCAAAGCCGAGCCCGACAGCC	2406
Q	y	1581	ACCTATGGACCTTCAACTCTCTGGGGCGGCGCTGATGATCCCTTAATACAGGTATCAGC	1640	Q	y	2661	ATGATCTCTCAACTGTGGAGGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCT	2720
D	b	1330	ACCTATGGACCTTCAACTCTCTGGGGCGGCGCTGATGATCCCTTAATACAGGTATCAGC	1389	D	b	2407	ATGATCTCTCAACTGTGGAGGCGCGGCACTTCCCAACGGCAACCTCAGCCAGCTGGCT	2466
Q	y	1641	CTCCTCATCCCGCCAGATGCCATACCCGAGGGAGATCTATGAGATCTACTCAGCGTG	1700	Q	y	2721	GCAGCAGTGGCTGAGCTGGGCGAGCGAGCTGGGCTCTTACAGTGTTCGAGGGCTGAG	2780
D	b	1390	CTCCTCATCCCGCCAGATGCCATACCCGAGGGAGATCTATGAGATCTACTCAGCGTG	1449	D	b	2467	GCAGCAGTGGCTGAGCTGGGCGAGCGAGCGTGGGCTCTTACAGTGTTCGAGGGCTGAG	2526
Q	y	1701	CACAAGCGGAGAGCTGAGGTGGCTTACTGCTGTGAGCCCTGCTGAGTCCCATC	1760	Q	y	2781	TGCTGAGGCGGCGGCGAGCGGCGAGCACTTCACTCTCAACAGTGTGGGACCCCGACAGGAC	2840
D	b	1450	CACAAGCGGAGAGCTGAGGTGGCTTACTGCTGTGAGCCCTGCTGAGTCCCATC	1509	D	b	2527	TGCTGAGGCGGCGGCGAGCGGCGAGCACTTACACTCTCACAGCTTTGGCACCCCAAGGAC	2586
Q	y	1761	GTTAGCTGTGACCCCTCGGCTGTGCTGACCCGCGGCGCTGATGATCCCTTAATACAGGTATCAGC	1820	Q	y	2841	AGGAGAGAGCGGAGCAGGGGCGCTTCCCGACACCGGGGAGA	2881
D	b	1510	GTTAGCTGTGACCCCTCGGCTGTGCTGACCCGCGGCGCTGATGATCCCTTAATACAGGTATCAGC	1569	D	b	2587	AGGAGAGAGCGGAGCAGGGGCGCTTCCCGACACCGGGGAGA	2627
Q	y	1821	TGTGGGAGCGCCAGCCCTGACAGCTGGAGCTTCCGCTCAAAAAGCACTGTGGGAGGGC	1880					
D	b	1570	TGTGGGAGCGCCAGCCCTGACAGCTGGAGCTTCCGCTCAAAAAGCACTGTGGGAGGGC	1629					
Q	y	1881	AGCTGGAGCAGGATGTGCTGACCTGGCAGGAGGCGCCCTCCACTCTACTACTGTC	1940					
D	b	1630	AGCTGGG---AGGATGTGCTGACCTGGGAGGAGGCGCCCTCCACTCTACTACTGTC	1686					
Q	y	1941	CAGCTGGAGGCGCAGTGTGCTTACGCTTTCACCGAGCAGCTGGGCGGCTTGGCCCTGGTG	2000					

RESULT 5  
AAK52261  
ID AAK52261 standard; cDNA; 2907 BP.  
XX  
AC AAK52261;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 806.  
XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

FN WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00496914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.

XX 15-SEP-2000; 2000US-00863561.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

XX Ma Y, Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX Xue AU, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

XX P-PSDB; AAM79128.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful

XX in diagnosis and gene therapy.

XX Claim 1; Page 2691-2694; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to

XX cytokine, cell proliferation or cell differentiation or which may induce

XX production of other cytokines in other cell populations. The

XX polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX e.g. stem cell growth factor activity, haematopoiesis regulating

XX activity, tissue growth factor activity, immunomodulatory activity and

XX activin/inhibin activity and may be useful in the diagnosis and/or

XX treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX inflammation. Note: Records for SEQ ID NO. 2110 (AAK52581), 2111

XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

XX sequence listing were missing at the time of publication

XX Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;

XX Query Match 79.7%; Score 2296.2; DB 4; Length 2907;

XX Best Local Similarity 92.0%; Pred. No. 0;

XX Matches 2565; Conservative 0; Mismatches 38; Indels 186; Gaps 6;

QY 133 TCGCGCTTGGCTCCCGGCTCGGGTCCAGCAGAGTGCACCGCGCCACCCAGTGC 192

DB 170 TGGCGGCAGACAGCAGCAGTGGAGCGCCAGCAGAGTGCCACCGTGGCCACCCAGTGC 229

QY 193 CTGGTGCACACCGGACCTGCTTCCCACTTCTGGTGGAGCCCGAGGATGTATCATCG 252

DB 230 CTGGTGCACACCGGACCTGCTTCCCACTTCTGGTGGAGCCCGAGGATGTATCATCG 289

QY 253 TCAAGAACAGCAGTGTGCTGTGTGCAAGCCGTCGCCGACGACGATCTTCTTCA 312

DB 290 TCAAGAACAGCAGTGTGCTGTGTGCAAGCCGTCGCCGACGACGATCTTCTTCA 349

QY 313 AGTGCAACCGGGAGTGGGTGGCGCCAGGTGGACACACGTGATCGAGCGCAGCAGCGGA 372

DB 350 AGTGCAACCGGGAGTGGGTGGCGCCAGGTGGACACACGTGATCGAGCGCAGCAGCGGA 409

QY 373 GCAGTGGTGAGCCGACATGGAGGTCCGCAATTAATGTCTCAAGGAGCAGGTCCGAGG 432

DB 410 GCAGTGGGCTGCCACCACATGAGGTCCCAATTAATGTCTCAAGGAGCAGAGTCCGAGAAG 469

QY 433 TGTTCGGGCTGGAGGAATACTGGTGCAGTGGTGCATGGAGTCCCTCGGSCACCAACA 492

DB 470 TGTTCGGGCTGGAGGAATACTGGTGCAGTGGTGCATGGAGTCCCTCGGSCACCAACA 529

QY 493 AGAGTCAGAGGCTTACATCCGCATAGCCAGATTGCGCAAGAATCTTCAGCAGGAGCCGC 552

DB 530 AGAGTCAGAGGCTTACATCCGCATAGCCAGATTGCGCAAGAATCTTCAGCAGGAGCCGC 589

QY 553 TGGCCAAAGGAGTCTCCCTGGAGCAGGCGCATGCTGCTGCCCTCGGCTCACCGAGGSCA 612

DB 590 TGGCCAAAGGAGTCTCCCTGGAGCAGGCGCATGCTGCTGCCCTCGGCTCACCGAGGSCA 649

QY 613 TCCCTCCAGCCGAGGTGGAGTGGCTCCGGAAACAGAGGACCTGGTGGACCCGCTCCCTGGACC 672

DB 650 TCCCTCCAGCCGAGGTGGAGTGGCTCCGGAAACAGAGGACCTGGTGGACCCGCTCCCTGGACC 709

QY 673 CCAATGTATATACATCACCGGGAGCACAGCCTGGTGGTGGACAGGCGCCGCTTGTGACA 732

DB 710 CCAATGTATATACATCACCGGGAGCACAGCCTGGTGGTGGACAGGCGCCGCTTGTGACA 769

QY 733 CGGCCCAACTACCTGCTGCGCCCAAGAACATCTGTGGCAGCTGCGCCGAGCGCCTCGGCTG 792

DB 770 CGGCCCAACTACCTGCTGCGCCCAAGAACATCTGTGGCAGCTGCGCCGAGCGCCTCGGCTG 829

QY 793 CTGTATCTGTCTAC----- 806

DB 830 CTGTATCTGTCTACGGTGGGCCCCGGGACTCCCTGCTCACAGGAGAGGCACTCGGCTGC 889

QY 807 ----- 825

DB 890 CCTTGGSCAGTACATGTGGCTGTCTCTCTGTCCGGCCAGTGAACGGTGGTGGTGA 949

QY 826 CFTGGACCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 885

DB 950 CFTGGACCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1009

QY 886 GGAGTGTGACCAACCGGCGCTCTCAACGGGGCGCTTCTGTGAGGGGCGCAATGTGC 945

DB 1010 GGAGTGTGACCAACCGGCGCTCTCAACGGGGCGCTTCTGTGAGGGGCGCAATGTGC 1069

QY 946 ATGACCGACCGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1005

DB 1070 AGAA---AAGACGCTGCGCCACCTGTGCGCAGTGGACGCGCAGTGGAGCCCGTGGAGCA 1126

QY 1006 AGTGTGCGGCTGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1065

DB 1127 AGTGTGCGGCTGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1186

QY 1066 CACCCCGCAACCGAGGGAGGAGTGGCAGGGCACTGACCTGGACACCCGCAACTGTGTACCA 1125

DB 1187 CACCCCGCAACCGAGGGAGGAGTGGCAGGGCACTGACCTGGACACCCGCAACTGTGTACCA 1246

QY 1126 GTGACCTCTGTGTACACA----- 1143

DB 1247 GTGACCTCTGTGTACACA----- 1306

QY 1144 ---GTGCTTCTGGCCCTGAGGAGTGGGCTCTATGTGGGCTCATGCGCGTGGCGCTCT 1200

DB 1307 CAGTGTCTTCTGGCCCTGAGGAGTGGGCTCTATGTGGGCTCTATGTGGGCTCTATGTGGGCTCT 1366

QY 1201 GCCTGGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260

DB 1367 GCCTGGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1426

QY 1261 ACTCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320

DB 1427 ACTCAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1486

QY 1321 CCAGCAAGGAGCAGAACCCCATCTGTCTACCATCCAGCGGACCTGAG---CACCAACA 1377

Db	1487	CGAGCAAGAGAGACACCCCACTCTCTCACTATCCAGCGGACCTCAGCACACACCA	1546
QY	1378	CCACCTACACAGGCGAGTCTCTGTCCCGGAGGATGGGCCACGCCCAAGTTCACAGTCA	1437
Db	1547	CCACCTACACAGGCGAGTCTCTGTCCCGGAGGATGGGCCACGCCCAAGTTCACAGTCA	1606
QY	1438	CCAAATGGGACCTGTCTGACGCCCCCTGGGTGGCGCGCCACACACATGACACAGCTCTC	1497
Db	1607	CCAAATGGGACCTGTCTGACGCCCCCTGGGTGGCGCGCCACACACATGACACAGCTCTC	1666
QY	1498	CCACCTCTGAGGCGGAGGAGTCTCTCCCGCTCTCCACCCAGAACTACTTCGCTCCC	1557
Db	1667	CCACCTCTGAGGCGGAGGAGTCTCTCCCGCTCTCTCCACCCAGAACTACTTCGCTCCC	1726
QY	1558	TGCCCCGAGGACACACGAAATGACCTATGGGACCTTCAACTCTCTGGGGGCGGCTGA	1617
Db	1727	TGCCCCGAGGACACACGAAATGACCTATGGGACCTTCAACTCTCTGGGGGCGGCTGA	1786
QY	1618	TGATCCCTAATACAGGTATCAGCTCTCTATCCCGCCAGATGCCATACCCGGAGGAGA	1677
Db	1787	TGATCCCTAATACAGGTATCAGCTCTCTATCCCGCCAGATGCCATACCCGGAGGAGA	1846
QY	1678	TCTATGAGATCTACTCTCAGCTGCACAAAGCGGAGACGTGAGGTTGCCCTAGCTGGCT	1737
Db	1847	TCTATGAGATCTACTCTCAGCTGCACAAAGCGGAGACGTGAGGTTGCCCTAGCTGGCT	1887
QY	1738	GTACAGCCCTGCTGAGTCCATCGTTAGCTGTGACACCCCTGGGCTCTGCTACCCGGC	1797
Db	1888	-----AGTGTGACCCCTGGGCTCTGCTACCCGGC	1921
QY	1798	CAGTCATCTGGCTATGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGGCC	1857
Db	1922	CAGTCATCTGGCTATGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGGCC	1981
QY	1858	TCAAAAGAGCTGCTGGGAGGAGCTGGGAGCAGGATGTCTGCACCTGGGCGAGAGG	1917
Db	2038	TCAAAAGAGCTGCTGGGAGGAGCTGGGAGCAGGATGTCTGCACCTGGGCGAGAGG	2038
QY	1918	CGCCCTCCCACTCTACTACTGACAGCTGGAGGCGAGTGTCTGCTCTTCCACGAGC	1977
Db	2039	CGCCCTCCCACTCTACTACTGACAGCTGGAGGCGAGTGTCTGCTCTTCCACGAGC	2098
QY	1978	AGCTGGGCGCTTGGCCCTGTGGGAGGAGCCCTCAGCTGGCTGGCCCAAGCGCTCA	2037
Db	2099	AGCTGGGCGCTTGGCCCTGTGGGAGGAGCCCTCAGCTGGCTGGCCCAAGCGCTCA	2158
QY	2038	AGCTGCTTCTGTTGGCGCGGTGGCTGCACCTCCTCGAGTACAAATCCGGGTCTACT	2097
Db	2159	AGCTGCTTCTGTTGGCGCGGTGGCTGCACCTCCTCGAGTACAAATCCGGGTCTACT	2218
QY	2098	GCCTGCATGACACCCACGATCACTCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGG	2157
Db	2219	GCCTGCATGACACCCACGATCACTCAAGGAGGTGGTGCAGCTGGAGAGCAGCTGGGG	2278
QY	2158	GACAGCTATCCAGAGACCGGCTCTGCACTTCAAGGACAGTTACACACCTGGCC	2217
Db	2279	GACAGCTATCCAGAGACCGGCTCTGCACTTCAAGGACAGTTACACACCTGGCC	2338
QY	2218	TATCATCCAGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCTGTGAGTACCAAG	2277
Db	2339	TATCATCCAGATGTGCCAGCTCCCTGTGGAAGAGTAAGCTCTGTGAGTACCAAG	2398
QY	2278	AGATCCCTTTTATCACAATCTGGAATGGACGCGAGGGTACTTGACATGCACCTCACCC	2337
Db	2399	AGATCCCTTTTATCACAATCTGGAATGGACGCGAGGGTACTTGACATGCACCTCACCC	2458
QY	2338	TGAGCGTCTAGCCCGACCTAGTGTGACCTGGCTGCAAGCTGGGTGGTGGCAGGTGG	2397
Db	2459	TGAGCGTCTAGCCCGACCTAGTGTGACCTGGCTGCAAGCTGGGTGGTGGCAGGTGG	2518
QY	2398	AGGGCGAGCGGAGAGCTTTCAGCATCACTTCAACATCACCAAGGACACAAAGTTTGCTG	2457
Db	2519	AGGGCGAGCGGAGAGCTTTCAGCATCACTTCAACATCACCAAGGACACAAAGTTTGCTG	2578
QY	2458	AGCTGCTGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCA	2517
Db	2579	AGCTGCTGCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCA	2638
QY	2518	AGATCCCTTCTCTCATTCGGCAGAAATAATTTCCAGCTGGACCCACCTGTAGCGGG	2577
Db	2639	AGATCCCTTCTCTCATTCGGCAGAAATAATTTCCAGCTGGACCCACCTGTAGCGGG	2698
QY	2578	GTGCCAGCTGGCGGACTGTGGCCCCAGAACTCCACTGGACAGCCATCTCAGCTTCTTG	2637
Db	2699	GTGCCAGCTGGCGGACTGTGGCCCCAGAACTCCACTGGACAGCCATCTCAGCTTCTTG	2758
QY	2638	CTTCCAGGCCAGCCCCACAGCCATCTCTCAACTGTGGGAGGCGGCACTTCCCCCA	2697
Db	2759	CTTCCAGGCCAGCCCCACAGCCATCTCTCAACTGTGGGAGGCGGCACTTCCCCCA	2818
QY	2698	ACGGCAACCTCAGCAGCTGGCTGCAGCAGTGGCTGGACCTGGGCCAGAGCTGGCC	2757
Db	2819	ACGGCAACCTCAGCAGCTGGCTGCAGCAGTGGCTGGACCTGGGCCAGAGCTGGCC	2878
QY	2758	TCTTCAAGTGTGGAGGCTGAGTGTGA	2786
Db	2879	TCTTCAAGTGTGGAGGCTGAGTGTGA	2907
RESULT 6			
AAV52940			
ID	AAV52940 standard; cDNA; 3014 BP.		
XX	AAV52940;		
AC	XX		
XX	XX		
DT	25-MAR-2003 (revised)		
DT	21-DEC-1998 (first entry)		
XX	Rat UNC-5 homologue unc5h-1 cDNA.		
DE	XX		
XX	UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;		
KW	diagnosis; therapy; ds.		
KW	Rattus sp.		
XX	XX		
OS	XX		
Key	Location/Qualifiers		
PH	1..2697		
FT	/*tag= a		
FT	XX		
XX	WO9837085-A1.		
PN	XX		
XX	27-AUG-1998.		
PD	XX		
PD	XX		
XX	19-FEB-1998; 98WO-US003143.		
PF	XX		
XX	19-FEB-1997; 97US-00808982.		
PR	XX		
XX	(REGC ) UNIV CALIFORNIA.		
PA	XX		
XX	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;		
PI	XX		
XX	WPI; 1998-495364/42.		
DR	P-PSDB; AAW78898.		
DR	XX		
XX	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and		
PT	the biopharmaceutical industry.		
XX	XX		
XX	Claim 7; Page 15-17; 32pp; English.		
PS	XX		
CC	This cDNA, termed unc5h-1, comprises a rat homologue of Caenorhabditis		
CC	elegans unc-5. Rat unc5h-1 and unc5h-2 (see AAV52942) cDNAs were isolated		
CC	from an E18 brain cDNA library. The predicted proteins (see AAW78898 and		
CC	AAW78900) show similarity with UNC-5. They are predicted to be involved		
CC	in cell migration and axon guidance, and are characterised as receptor		
CC	proteins for netrins. Gene expression is observed in regions where		
CC	differentiating neurons are undergoing axogenesis. Human unc5h-1 (see		

CC AAV52941) and unc5h-2 (see AAV52943) cDNAs are also claimed. Vertebrate  
CC UNC-5 proteins may be produced recombinantly from transfected host cells  
CC by utilising these vertebrate UNC-5 nucleic acids. The invention also  
CC provides unc-5 hybridisation probes and primers, vertebrate UNC-5-  
CC specific binding agents such as specific antibodies, and methods of  
CC making and using the subject compositions in diagnosis (e.g. genetic  
CC hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g.  
CC gene therapy to modulate vertebrate unc-5 gene expression) and in the  
CC biopharmaceutical industry (e.g. as immunogens, reagents for modulating  
CC cell guidance, reagents for screening chemical libraries for lead  
CC pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI  
CC field.)  
XX

SQ Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;  
Query Match 77.6%; Score 2235.8; DB 2; Length 3014;  
Best Local Similarity 88.5%; Pred No. 0;  
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;  
QY 87 ATGGCCCTCGGCGCGGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 146  
DB 1 ATGGCCCTCGGCGCGGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTTGGCTT 60  
QY 147 CGCGGCTCGGCTCGGCGGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTTGGCT 206  
DB 61 CGTGGTTGGGTCGCCAGCAGTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTTGGCT 120  
QY 207 GACCTGCTTCCCTCCCTCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCG 266  
DB 121 GACCTGCTTCCCTCCCTCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCG 180  
QY 267 GTGCTGCTTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 326  
DB 181 GTGCTGCTTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 240  
QY 327 TGGGTCGCGGCTGCGGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 386  
DB 241 TGGGTCGCGGCTGCGGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 300  
QY 387 ACCATGGAGTCCGCTTGAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 446  
DB 301 ACCATGGAGTCCGCTTGAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 360  
QY 447 GAATATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 506  
DB 361 GAATATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 420  
QY 507 TACATCGGATGCGGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 566  
DB 421 TACATCGGATGCGGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 480  
QY 567 TCCCTGGAGCAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 626  
DB 481 TCCCTGGAGCAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 540  
QY 627 GTGGAGTGTCTCGGAAACGAGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCG 686  
DB 541 GTGGAGTGTCTCGGAAACGAGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCG 600  
QY 687 ACAGCGGAGCAGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 746  
DB 601 ACAGCGGAGCAGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 660  
QY 747 TGGCTGGCCAGAAATCGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 806  
DB 661 TGGCTGGCCAGAAATCGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 720  
QY 807 GTGAACGCTGGTGTGTGTCGAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 866  
DB 721 GTGAACGCTGGTGTGTGTCGAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 780  
QY 867 GCTTGGCAGAAACGAGCTGTGGCCAGTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 926

DB 781 GGCTGGCAGAAACGAGCCGAGCTGCAACCAACCCGCGCACCTCTCAACGGGGCGCCCTTC 840  
QY 927 TGTGAGGGGCGAAGATGTCATGACCGCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 986  
DB 841 TGTGAGGGGCGAAGATGTCATGACCGCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 897  
QY 987 AGCTGGAGCCCGCTGGAGCAAGTGTGGCCCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1046  
DB 998 AGCTGGAGTTCGTGGAGTAAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 957  
QY 1047 CGTGAAGTCTCTGACCCAGCACCCCGCAACGAGGGGAGGAGTCCAGGCGACTGACCTG 1106  
DB 958 CGCAGTGTCTGACCCAGCACCCCGCAANTGAGGTGAGGAGTGTGGGGTGTGACCTG 1017  
QY 1107 GACACCCGCAACTGTACAGTGAACCTCTGTGTGTACACAGTGTCTTGGCCCTCTGAGGACCTG 1166  
DB 1018 GACACCCGCAACTGTACAGTGAACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1077  
QY 1167 GCCCTCTATGTGGCCCTCATGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1226  
DB 1078 GCTCTCTACATGGGCTTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1137  
QY 1227 ATCTCTCTTTATTTGCCGAAAGAGGAGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1286  
DB 1138 GGACTCATTTACTGTGCGAAGAGGAGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1197  
QY 1287 CTCACTTCAGGCTTCAGCCCGCTCAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1346  
DB 1198 CTCACTTCAGGCTTCAGCCCGCTCAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1257  
QY 1347 CTCACTTCAGGCTTCAGCCCGCTCAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCTG 1403  
DB 1258 CTCACTTCAGGCTTCAGCCCGCTCAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCTG 1317  
QY 1404 CGCAGGATGGGCGCCAGCCCAAGTTCCAGCTCACAATGGGCAAGCTGTCTGAGGCTGTCTGAGGCTGT 1463  
DB 1318 AGCAGGATGGAGCCAGCCCGCTCAGCATCAAGCTCACAATGGGCAAGCTGTCTGAGGCTGTCTGAGGCT 1377  
QY 1464 GFTGGCGGCGCCACACACTGACCAAGCTCTCCACCTCTGAGGCGGAGGAGTGTCTGCT 1523  
DB 1378 GGGAGTGGCGCGCATACGTTGACCAAGCTCTCCACCTCTGAGGCTGTCTGAGGCTGTCTGAGGCT 1437  
QY 1524 TCCCGCTCTCCACCGAGCAACTTCTCCGCTCTCCGCTCTCCGCTCTCCGCTCTCCGCTCTCCGCTCTCCGCT 1583  
DB 1438 TCCCGCTCTCCACCGAGCAACTTCTCCGCTCTCCGCTCTCCGCTCTCCGCTCTCCGCTCTCCGCTCTCCGCT 1497  
QY 1584 TATGGACCTTCAACTTCTCCGCGGCGGCTGTGATGATCCCTAAATACAGGATACAGCTCT 1643  
DB 1498 TACGGGACCTTCAACTTCTCCGCGGCGGCTGTGATGATCCCTAAATACAGGATACAGCTCT 1557  
QY 1644 CTCACTCCCGGAGTGCCTATCCCGAGGGAAGATCTATGAGATCTATGAGATCTATGAGATCTATGAGATCT 1703  
DB 1558 CTCACTCCCGGAGTGCCTATCCCGAGGGAAGATCTATGAGATCTATGAGATCTATGAGATCTATGAGATCT 1617  
QY 1704 AAGCGGAGAGCTGAGGTTGCCCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1763  
DB 1618 AAGCGGAGAGCTGAGGTTGCCCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1677  
QY 1764 AGCTGTGAGACCCCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1823  
DB 1678 AGCTGTGAGACCCCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1737  
QY 1824 GGGAGCCAGCCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1883  
DB 1738 GGGAGCCAGCCCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1797  
QY 1884 TGGGAGCAGGATGTCTGCACTGTGGCGAGGAGGCGCTCCCACTCTACTACTACTACTACTACTACTACTACT 1943  
DB 1798 TGGG---AGGATGTGCTGTGCACTGTGGCGAGGAGTCACTTCCCACTCTACTACTACTACTACTACTACTACT 1854  
QY 1944 CTGAGGCGGAGTGTCTGCTTACCGAGGAGTGTGGCGGCTTGTGGCGGCTTGTGGCGGCTTGTGGCGGCTTGT 2003  
DB 1855 CTGAGGCGGAGTGTCTGCTTACCGAGGAGTGTGGCGGCTTGTGGCGGCTTGTGGCGGCTTGTGGCGGCTTGT 1914

local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1; calcium-independent alpha-latrotoxin receptor; CIRL; trke; synapsin 1A; epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic; ss; tumour necrosis factor alpha; TNF-alpha; rat.

Rattus sp.

OS OS

Key Location/Qualifiers

CDS 1..2697

FT /tag= a

FT /product= "Rat netrin receptor UNC5H1"

XX W0200175440-A2.

XX PD 11-OCT-2001.

XX PF 02-APR-2001; 2001WO-GB001486.

XX PR 31-MAR-2000; 2000GB-00007880.

XX PR 26-MAY-2000; 2000GB-00012768.

XX PA (WELP-) WELFIDE CORP.

XX PI Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;

XX PI WPI; 2002-010813/01.

XX DR P-PSDB; AAU10543.

XX DR Novel chronic animal model of schizophrenia, useful for identifying anti-psychotic drugs and genes that are associated with schizophrenia.

XX PS Claim 1; Fig 8a; 79pp; English.

XX CC The invention relates to YSG polynucleotide fragments for use in diagnosing and/or developing treatments for schizophrenia using chronic animal models. The polynucleotides and their encoded polypeptides are used for identification of compounds which modulate the expression of YSG molecules, leading to the manufacture of schizophrenia medications. The sequences can also be used for testing candidate compounds for any effect on the polypeptides. Anti-schizophrenic effects of a compound can be determined by measuring local cerebral glucose utilisation (LCGU) or comparing its expression level with that of a control group. The sequences are useful in the identification of genes associated with schizophrenic states and in the development of an antibody. The sequences of the invention include phosphodiesterase 1-alpha, calcium-independent alpha-latrotoxin receptors (CIRL)-1,2&3 epithelial discoidin domain receptor 1 (trke), netrin receptor (UNC5H1), synapsins 1A and 1B and tumour necrosis factor (TNF) alpha. This sequence represents rat netrin receptor UNC5H1 (YSG7) DNA

XX SQ Sequence 2697 BP; 541 A; 864 C; 766 G; 526 T; 0 U; 0 Other;

Query Match 76.0%; Score 2189.4; DB 6; Length 2697;

Best Local Similarity 88.9%; Pred. No. 0;

Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;

Qy 87 ATGCGCGTCCGCGCGCGCTGTGGCAGCGCTCTCTGGGCATAGTCTCGCGCTTGGCTC 146

Db 1 ATGGCCGTCCGCGCGCGCTGTGGCAGTGCTCTCTGGGCATAGTCTCGCGCTTGGCTT 60

Qy 147 CGCGCTCGGCTGCCAGCAGAGTGCCACCGTGGCCAAACCCAGTGCCTGTGGCCACCG 206

Db 61 CGTGGTTCGGGTGCCAGCAGAGTGCCACCGTGGCCAAATCCAGTGCCTGGCCACCGC 120

Qy 207 GACCTGCTTCCCACTTCTCTGGAGCCCGAGGATGTACATCGTCAAGAAACAAGCA 266

Db 121 GACCTGCTGCCCACTTCTCTGGAGCCTTGGAGACGTGTACATTTGTCAAGAACAGCG 180

Qy 267 GTGCTGCTTGTGTGCAAGCGCGTCCCGCCAGCAGATCTTCTTCAAGTCAACGGGAG 326

Db 181 GTGTGTTGTGTGTGCAAGCGTGTGCTGCCACCCAGATCTTCTTCAAGTCAATGGGAA 240

Qy 327 TGGGTGCCCGAGGTGGACACGTGATCGAGCGCAGACAGCGGAGCAGTGGTGGCGG 386

2004 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTTGTGGCGCGTGGCC 2063

Db 1915 GAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAGGCTCCTTCTGTTGTCTCCGTTGGCC 1974

Qy 2064 TGCACCTCCTCGAGTACAACATCCGGGTCTACTGCTGCATACACCCAGATGCACCTC 2123

Db 1975 TGTACGTCCCTTGAGTACAACTCCGAGTGTACTGCTTACACACCCAGCAGCTCTC 2034

Qy 2124 AAGGAGTGTGTGAGTGGAGAGAGCTGGGGGACAGCTGATCCAGGAGCCACCGGTC 2183

Db 2035 AAGGAGTGTGTGAGTGGAGAGAGCTAGTGTGACAGCTGATCCAGGAGCCTCGGTC 2094

Qy 2184 CTCACCTTCAAGAGATACCAACACCTCGGCTATFCAATCAAGATGTGCCAGTCC 2243

Db 2095 CTGCACCTCAAGAGATACCAACACCTTACCTTCCATCCAGCAGCTGCCAGCTCC 2154

Qy 2244 CTGTGAAGAGTAACTCTTGTGAGTACAGCTACAGAGATCCCTTTTATCACTGTGAAT 2303

Db 2155 CTGTGAAGAGAGAGCTTGTGAGTACAGCTACAGAGATCCCTTTTACCACTGTGAAC 2214

Qy 2304 GGCAACGAGCGGTACTTGTGCACTGCACCTTCAACCTGGAGGCTGTCAAGCCAGCACTAGT 2363

Db 2215 GGCAACGAGCAGTATCTGCACTGCACCTTCAACCTGGAGGCTCAACGCGCAGCAGC 2274

Qy 2364 GACCTGCGCTGCAAGCTGTGGGTGTGGAGTGGAGGCGAGCGGAGAGCTTCAAGATC 2423

Db 2275 GACCTGCGCTGCAAGTGTGGGTGTGGAGTGGAGGAGTGGGAGAGCTTCAACATC 2334

Qy 2424 AACTTCAACATCAACCAAGGACACAAGTTTGTCTGAGCTGTGGTCTGTGGAGTGAAGCG 2483

Db 2335 AACTTCAACATCACTAAGGACACAAGTTTGTCTGAAATTTGGCTCTGGAGTGAAGGG 2394

Qy 2484 GGGTCCCGAGCCCTGGTGGGCCAGTCCCTTCAAGATCCCTTCTCATTTGGCGCAAG 2543

Db 2395 GGGTCCCGAGCCCTGGTGGGCCAGTCCCTTCAAGATCCCTTCTCATTTGGCGCAAG 2454

Qy 2544 ATAAATTCAGCTGGACCCACCTGTAGTGGCGGCTGGAGTGGGAGCTTGGCCGAG 2603

Db 2455 ATCATCGCAGTCTGGACCCACCTTGCAGCGCGGCGCGAGTGGAGAACTCTAGCCGAG 2514

Qy 2604 AAATCCAGCTGGACGCACTCTGAGCTTCTTGTCTTCCAGCCAGCCAGCCAGCCAGCATG 2663

Db 2515 AAATCTCACCTGGACGCACTCTGAGCTTCTTGTCTTCCAGCCAGCCAGCCAGCATG 2574

Qy 2664 ATCTCTCAACCTGTGGAGCGCGGCACTTCCCAACGCAACCTCAGCCAGCTGGCTGCA 2723

Db 2575 ATCTCTCAACCTATGGAGGCGCGGCACTTCCCAACGCAACCTCAGCCAGCTGGCAGCA 2634

Qy 2724 GCAGTGGCTGAGTGGCGCGGCGAGCAGCTGGCTTCTTCAAGTGTGGAGGCTGAGTGC 2783

Db 2635 GCTGTGGCGGAGTGGCGCGGCGAGATGCTGGCTTCTTCAAGTGTGGAGGCGGAGTGT 2694

Qy 2784 TGAGCGCGCGCAGGCGCGG -ACACCTACACTCTACAGCTTTGGCAGCCACCCAGGAGCA 2841

Db 2695 TGAGACCGAGCGCGGCGGATGCTTACCTTACAGCTTTGACACTTGCACCTGCCAGGACA 2754

Qy 2842 GGCAGAGCGGAGAGCGGCGGCTTCCCAACACCGGGG 2878

Db 2755 GGCAGAGCGGAGAGCGGCGGCTTCCCAACACCGGGG 2791

RESULT 7

AA516843

ID AA516843 standard; cDNA; 2697 BP.

XX AA516843;

AC AA516843;

XX 14-FEB-2002 (first entry)

DT Rat netrin receptor UNC5H1 (YSG7) cDNA.

XX YSG; YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;

Db 241 TGGGTCCGCCAGTTCGATCACTAAATTAACCGACGACCGGACGACGAGCGGATGGCA 300  
QY 387 ACCATGAGGTCCGCAATTAATGTCTCAAGGACGAGGTGAGAGGTGTTCCGGCTGGAG 446  
Db 301 ACCATGAGGTCCGCAATTAATGTCTCAAGGACGAGGTGAGAGGTGTTCCGGCTGGAG 360  
QY 447 GAATACCTGGTCCAGTGGTGGCATGAGCTTCCTCGGSCACCAACAGAGTCAAGAGCC 506  
Db 361 GAATACCTGGTCCAGTGGTGGCATGAGCTTCCTCGGSCACCAACAGAGTCAAGAGCC 420  
QY 507 TACATCCGATAGCCAGATTGGCGCAAGACTTCGAGCAGGAGCGCTGGCCAGAGGTG 566  
Db 421 TACATCCGATAGCCAGATTGGCGCAAGACTTCGAGCAGGAGCGCTGGCCAGAGGTG 480  
QY 567 TCCCTGAGCAGGSCATCGTCTCCCTCGGTCACCGGAGGACATCCCTCAGCCGAG 626  
Db 481 TCACTGAGCAAGGCAATGTACTACCTTGTGCCCCCAGAGGAATCCCCCCAGCTGAG 540  
QY 627 GTGAGTGGCTCCGGAACGAGGACCTGTGGACCCGCTCCCTGGACCCCAATGTATACATC 686  
Db 541 GTGAGTGGCTTCGAATGAGGACCTGTGGACCCCTCCCTCGATCCCAATGTGTATC 600  
QY 687 ACCGGGAGCACAGCTGTGTGGTGGAGACGCGCCCTTGTGACACGGCCAACTACACC 746  
Db 601 ACCGGGAGCACAGCTGTGTGGTGGAGACGCGCCCTTGTGACACGGCCAACTACACC 660  
QY 747 TGGTGGCCAAAGAACATCGTGGACCGTGGCGGAGCGCCCTCGCTGTGTATCGCTAC 806  
Db 661 TGTGTGGCCNAGAACATCGTGGACCGTGGCGGAGCACCTCTCAGCGGTCAATGTTTAT 720  
QY 807 GTGAACGGTGGTGGTGGACCGAGTGGTTCGTCGACGCGCAGCTGTGGSCGC 866  
Db 721 GTGAACGGTGGTGGTGGACCGAGTGGTTCGTCGACGCGCAGCTGTGGSCGT 780  
QY 867 GGTGTGCAGAAACGGAGCCGAGCTGCAACCAACCGCGCGCTCTCAACGGGGCGCTTTC 926  
Db 781 GGTGTGCAGAAACGGAGCCGAGCTGCAACCAACCGCGCACCTCTCAACGGGGCGCGCTTC 840  
QY 927 TGTGAGGGGAGAAATGTCATGACCGACCGTCTCCTCTCTCTGTGTCTCTGTGAGCGC 986  
Db 841 TGTGAGGGGAGAAATGTCAGAAACAGCGCTGGCCACCTCTG--TGCCAGTGGATGGG 897  
QY 987 AGCTGAGCGCGTGGAGCAAGTGGTGGCGTGGCTGGACTGACCCACCTGGCGGAGC 1046  
Db 898 AGCTGAGTTCGTGGAGTAAGTGGTCAAGCTGTGGCTTGAATGACCCACTGGCGGAGC 957  
QY 1047 CGTGAAGTCTGACCCAGCAACCCGGAACGGAGGGAGGTGCCAGGSCACTGACCTG 1106  
Db 958 CGCGAGTGTCTGACCCAGCAACCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTG 1017  
QY 1107 GACACCCGCAACTGTACAGTGAACCTGTGTACAGAGTGTCTGTGGCTGGAGGAGCTG 1166  
Db 1018 GACACCCGCAACTGTACAGTGAACCTGTGTACAGAGTGTCTGTGGCTGGAGGAGCTG 1077  
QY 1167 GGCCTATGTGGGCTCATCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGCTC 1226  
Db 1078 GCTCTCTACATCGGCTGTGCTGTGGTGTGTGCTCTTCTGTGTGTGTGTGTGTGTGTGT 1137  
QY 1227 ATCTCTGTTTATTCGGAAGAGAGGGGCTGGACTCAGATGTGGCTGACTGTGCTCAT 1286  
Db 1138 GGACTCATTTACTGTGCGAAGAGAGGGCTGGACTCAGATGTGGCTGACTGTGCTCAT 1197  
QY 1287 CTCACCTCAGGCTTCAGCCCGTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 1346  
Db 1198 CTCACCTCAGGCTTCAGGCTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 1257  
QY 1347 CTCACCATCAGCGGCTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 1403  
Db 1258 CTCACCATCAGCGGCTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 1317  
QY 1404 CGCGAGGATGGGCGGAGCCGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 1463

Db 1318 AGGAGGATGGACCCAGCCCAAGTTCAGCTCTCTAATGCTACCTGTCTCAGGCCACTG 1377  
QY 1464 GGTGGCGCGGCCACACACCTGCAACAGCTTCTCCACCTCTGAGGCGGAGAGTTCGTC 1523  
Db 1378 GGGAGTGGCGGCCATACGTTTGCAACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTC 1437  
QY 1524 TCCGCTCTCCACCCAGAACTACTTCGCTCCCTGCTGCCCCGAGGACACGCAACATGAC 1583  
Db 1438 TCCGCTCTCCACCCAAATCTCTTCGTTCCCTGCTGCCCCGAGGACACGCAACATGACC 1497  
QY 1584 TATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGGCTC 1643  
Db 1498 TAGGGACCTTCACTTCTCGGGGCGGCTGATGATCCCTAATACGGGATCAGGCTC 1557  
QY 1644 CTATCCCCCAGATGCAATACCCCGAGGAAAGATCTATCAGATCTACCTCAGCTGCAC 1703  
Db 1558 CTCATACCCCGGATGCCATCCCGAGGAAAGATCTACGAGATCTACCTCAGCTGCAC 1617  
QY 1704 AAGCGGAAGACGTGAGTTTGGCCCTAGCTGGCTGTGAGACCTCTGTGAGTCCCATGTT 1763  
Db 1618 AAGCAGAAACGTGAGTTTGGCCCTAGCTGGCTGTGAGACCTCTGTGAGTCCCATGTT 1677  
QY 1764 AGCTGTGGACCCCTCGGCTCTGCTCACCGGCCAGTCACTCTGGCTATGAGACCTGTT 1823  
Db 1678 AGCTGTGGGCCCCAGAGTCTCTGCTCACCGGCCAGTCACTCTGCAATGGACCTGTT 1737  
QY 1824 GGGAGCCACGCTCTGACAGTGGAGCGCTCGCTCAAAAAGCAGTGTGCGAGGCGAGC 1883  
Db 1738 GGAGAGCCAGCCCTGACAGTGGAGTCTGGCTCTCAAAAAGCAGTGTGCGAGGCGAGT 1797  
QY 1884 TGGGAGCAGATGTGCTGGACCTGGGCGGAGGAGCGCTCCACCTCTACTACTGCTGCG 1943  
Db 1798 TGGG---AGATGTGCTGACCTTGTGAGGAGTCACTTCCACCTCTACTACTGCTGCG 1854  
QY 1944 CTGGAGCCAGTGCCTGCTACGCTTTCAGCGAGCTGGGCGGCTTGGCCCTGTGGGA 2003  
Db 1855 CTGGAGCCGGGCTCTGCTATGCTTTCAGGAGCAGTGGGCGGCTTGGCCCTGTGGGA 1914  
QY 2004 GAGGCCCTCAGCTGCTGCGCCAAAGCGCTCAAAGTGTCTGTTTGGCGCGGTGGCC 2063  
Db 1915 GAGGCCCTCAGCTGCTGCGCCAAAGCGCTCAAAGTGTCTGTTTGGCGCGGTGGCC 1974  
QY 2064 TGACCTCTCGGTGACATCCGGCTTACTGCTGTGATGACACCAAGAGTCACTC 2123  
Db 1975 TGTACCTCTGAGTACCAATCCGAGTGTACTGCTTACAGACACCCACGAGCTCTC 2034  
QY 2124 AAGGAGTGGTGCAGCTGGAGAGCAGCTGGGAGGAGCTGATCCAGAGCCACGAGTC 2183  
Db 2035 AAGGAGTGGTGCAGCTGGAGAGCAGCTGAGTGGAGCAGCTGATCCAGAGCTCTCGCTC 2094  
QY 2184 CTGCACTTCAAGACAGTTTACCAACCTGGGCTTATCCATCCAGATGTGCCAGCTCC 2243  
Db 2095 CTGCACTTCAAGACAGTTTACCAACCTGCTCTCTCCATCCAGAGCTGCCAGCTCC 2154  
QY 2244 CTGTGAAGAGTAACTCTCTGCTGCTACAGGAGATCCCTTTTATCAGTCTGGAAT 2303  
Db 2155 CTGTGAAGAGCAGCTACTTGTGAGTACAGGAGATCCCTTTTACCAATCTGGAAC 2214  
QY 2304 GGACGCGAGCGGTACTTGCATCTGCACCTTCACTTGGAGCGTGTGAGCCGAGCTAGT 2363  
Db 2215 GGACCCAGCAGTATCTGCATCTGCACCTTCACTTGGAGCGGATCAAGCCAGCAGCAG 2274  
QY 2364 GACCTGGCTTCAAGCTGTGGTGTGGCAGGTGGAGGGGCGAGCGGAGAGCTTCAGCATC 2423  
Db 2275 GACCTGGCTTCAAGCTGTGGTGTGGCAGGTGGAGGGGAGATGGGAGAGCTTCAACATC 2334  
QY 2424 AACTTCAACATCAGCAAGAGCAGAGTGTGCTGAGCTGCTGGCTCTGGAGAGTGAAGC 2483  
Db 2335 AACTTCAACATCAGCAAGAGCAGAGTGTGCTGAAATTTGTGGCTCTGGAGAGTGAAGG 2394  
QY 2484 GGGGCTCCAGCGCTGTGGGCCCCAGTGTCTTCAAGATCCCTTCTCTTCTTCTTCTTCT 2543  
Db 2395 GGGGCTCCAGCGCTGTGGGCCCCAGTGTCTTCAAGATCCCTTCTTCTTCTTCTTCTTCT 2454







QY	1830	CCGAGCCCTGACAGCTGGAGCGCTGCGCTCAAAAAGACGTCGTCCGAGGGCAGCTGGGAG	1889
Db	720	CCCAGCCCTGACAGCTGGAGCGCTGGCCCTCAAAAAGACGTCGTCCGAGGG-AGCTGGGAG	778
QY	1890	CAGGATGTGCTGCACCTGGGCGAGGAGCGCCCTCCACACCTCTACTACTGCACGCTGGAG	1949
Db	779	GATG---TCTGCACTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCACGCTGGAG	834
QY	1950	GCCAGTGCCTGACTAGCTTTCAACGAGAGCTGGGCGCGCTTTGCCCTGGTGGGAGAGGCC	2009
Db	835	GCCAGTGCCTGCTAGCTTTCAACGAGAGCTGGGCGCGCTTTGCCCTGGTGGGAGAGGCC	894
QY	2010	CTCAGCGTGGCTGCGCCCAAGCGCCTCAAGCTGCTCTGTTTGGCGCGGTGGCGCTGCACC	2069
Db	895	CTCAGCGTGGCTGCGCCCAAGCGCCTCAAGCTGCTCTGTTTGGCGCGGTGGCGCTGCACC	954
QY	2070	TCCCTCGAGTACAACATCCGGGCTACTGCTGTCATGACACCCACGATGCACTCAAGGAG	2129
Db	955	TCCCTCGAGTACAACATCCGGGCTACTGCTGTCATGACACCCACGATGCACTCAAGGAG	1014
QY	2130	GTGGTGAGCTGGAGAGCAGCTGGGGGGAAGCTGATCCAGHAGCCACCGGTCCTGTCAC	2189
Db	1015	GTGGTGAGCTGGAGAGCAGCTGGGGGGAAGCTGATCCAGHAGCCACCGGTCCTGTCAC	1074
QY	2190	TTCAAGGACAGTTTACCACAAACCTTCGGCGCTATCCATCCACGATGTCGCCAGCTCCCTGTGG	2249
Db	1075	TT-AAGGACAGTTTACCACAACT--GCCCTATCATCCACGATGTCGCCAGCTCCCTGTGG	1131
QY	2250	AAGAGTAAGCTCCTGTTCAGCTACAGGAGATGCCCTTTTATCACTCTGGAATGGCAGC	2309
Db	1132	AAGAGTAAGCTCCTGTTCAGCTACAGGAGATGCCCTTTTATCACTCTGGAATGGCAGC	1191
QY	2310	CAGCGGTACTTGCATGCACTTCACCTTCGAGCGGTGCAGCCCGACGACTAGTGACCTG	2369
Db	1192	CAGCGGTACTTGCATGCACTTCACCTTCGAGCGGTGCAGCCCGACGACTAGTGACCTG	1251
QY	2370	GCTTGAAGCTGTGGGTGTGGCAGGTGGAGGGCGGACGGGCAGAGCTTTCAGCATCAACTTC	2429
Db	1252	GCTTGAAGCTGTGGGTGTGGCAGGTGGAGGGCGGACGGGCAGAGCTTTCAGCATCAACTTC	1311
QY	2430	AACATCACCAAGGACACAGGTTTGTGAGCTGCTGGCTCTGAGAGTGAGCGGGGTC	2489
Db	1312	AACATCACCAAGGACACAGGTTTGTGAGCTGCTGGCTCTGAGAGTGAGCGGGGTC	1371
QY	2490	CCAGCCCTGTGGGCGCCCGAGTGCCTTCAAGATCCCTTCCTCATTCGGCAGAGATAATT	2549
Db	1372	CCAGCCCTGTGGGCGCCCGAGTGCCTTCAAGATCCCTTCCTCATTCGGCAGAGATAATT	1431
QY	2550	TCCAGCTGGAACCCAGCTGTAGCGGGGTGCGGACTGCGGGACTTGCGCCAGAACTC	2609
Db	1432	TCCAGCTGGAACCCAGCTGTAGCGGGGTGCGGACTGCGGGACTTGCGCCAGAACTC	1491
QY	2610	CACCTGGAAGCCCATCTCAGCTTCTTTTGCTTCCAGCCCGAGCCCAAGCCATGATCCTC	2669
Db	1492	CACCTGGAAGCCCATCTCAGCTTCTTTTGCTTCCAGCCCGAGCCCAAGCCATGATCCTC	1551
QY	2670	AACCTGTGGAGCGCGGCACTTCCCAACGGCAACTCAGCCAGCTGGCTGCAGCAGTG	2729
Db	1552	AACCTGTGGAGCGCGGCACTTCCCAACGGCAACTCAGCCAGCTGGCTGCAGCAGTG	1611
QY	2730	GCTGGACTGGCCAGCCAGAGCTGGCCTC-TTCAGTG-TCGGAGGCTGAGTGTGAG	2787
Db	1612	GCTGGACTGGCCAGCCAGAGCGTGGCTTCTTTCAGTGTTTCGAGGCTGAGTGTGAG	1671
QY	2788	GCCGGCCAGCCCGACACCTTACACTCTCCACAGCTTTGGCA--CCCAACAGGACAGGCA	2845
Db	1672	GCCGGCCAGG--CGAACACTACAAATTTACAGTTTGGGAACCCCAAGGACAGGCA	1729
QY	2846	GAAGCCGGACAGGGGCCCTTCCCAACCGGGGAGA	2881
Db	1730	GAAGCCGGACAGGGGCTTTTCCAAAACCGGGGAGA	1765

RESULT 9  
AAC90958  
ID AAC90958 standard; cDNA; 9700 BP.  
XX AC AAC90958;  
XX  
XX  
DT 19-MAR-2001 (first entry)  
XX  
XX Plasmid pGC1037 nucleotide sequence SEQ ID NO:91.  
DE  
XX Human; *Caenorhabditis elegans*; UNC-5; splice variant; nematode worm;  
KW protein-protein interaction; identification; ss.  
XX  
XX Synthetic.  
OS  
XX WO2000073328-A2.  
FN  
XX 07-DEC-2000.  
PD  
XX 02-JUN-2000; 2000WO-EP005108.  
PF  
XX 01-JUN-1999; 99GB-00012755.  
PR  
XX (DEVG-) DEVGEN NV.  
PA  
XX Van Crieckinge W, Roelens I, Bogaert T, Verwaerde P;  
PI WPI; 2001-016508/02.  
DR  
XX Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and a  
PT human unc-5HS1 cDNA, useful in yeast two hybrid experiments for  
PT identifying unknown human cDNAs which encode proteins that interact with  
PT the human unc-5C protein.  
XX  
XX Example 4; Page 228-230; 246pp; English.  
PS  
XX The present invention describes 3 variants of human unc-5C cDNAs (unc-  
CC 5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-  
CC 5C transcripts, and a human unc-5HS1 cDNA which shares homology with the  
CC *Rattus norvegicus* unc-5HS1 cDNA. Also described are assays based on  
CC protein-protein-interactions between the unc-5 protein and a variety of  
CC different interacting proteins. The unc-5C variant cDNAs and unc-5HS1  
CC cDNA are useful in methods for identifying compounds which reduce or  
CC inhibit the lethal phenotype associated with the expression of the unc-5  
CC death domain in yeast. They are also useful in yeast two hybrid  
CC experiments for identifying unknown human cDNAs which encode proteins  
CC that interact with the human unc-5C protein. AAC90914 to AAC90971 and  
CC AAB50646 to AAB50693 represent sequences used in the exemplification of  
CC the present invention  
XX  
XX Sequence 9700 BP; 2604 A; 2193 C; 2236 G; 2666 T; 0 U; 1 Other;  
SQ

	Query Match	44.8%;	Score 1291.2;	DB 4;	Length 9700;
	Best Local Similarity	99.5%;	Pred. No. 1.8e-228;		
	Matches 1306;	Conservative	0;	Mismatches 3;	Indels 3; Gaps 1;
QY	1478	CACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGAGTTCGTTCCCGCCTCTCCAC	1537		
Db	7	CACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGAGTTCGTTCCCGCCTCTCCAC	66		
QY	1538	CCAGAACTACTTCCGTCCTCGTCCCGGAGGCACCGACCAATGACCTATGGGACCTTCAA	1597		
Db	67	CCAGAACTACTTCCGTCCTCGTCCCGGAGGCACCGACCAATGACCTATGGGACCTTCAA	126		
QY	1598	CTTTCCTCGGGGCGGCTCATGATCCCTAATACAGATATCAGCCTCTCATCCCCCAGA	1657		
Db	127	CTTTCCTCGGGGCGGCTCATGATCCCTAATACAGGAATCAGCCTCTCATCCCCCAGA	186		
QY	1658	TGCATATACCCCGAGGGAAGATCTTATGAGATCTACCTCATCGCTGTCACAAACCGGAAGACGT	1717		
Db	187	TGCATATACCCCGAGGGAAGATCTTATGAGATCTACCTCATCGCTGTCACAAACCGGAAGACGT	246		
QY	1718	GAGGTTGCCCTAGCTGGCTGTGCAGACCTGTGTAGTCCCATCTGTAGCTGTGGACCCCC	1777		

Db 247 GAGGTTGCCCTAGCTGGCTGTGAGTCCATGTTAGCTGTGACCCCC 306  
 QY 1778 TGGCGTCTCTGCTCACCAGGCGAGTCTCTGCTATGAGACCACTGTGGGAGCCAGCC 1837  
 Db 307 TGGGTCCTGCTCACCAGGCGAGTCTCTGCTATGAGACCACTGTGGGAGCCAGCC 366  
 QY 1838 TGACAGCTGAGCTGGCTGCTCAAAAGAGTGTGGGAGGCGAGCTGGGAGCAGATGT 1897  
 Db 367 TGACAGCTGAGCTGGCTGCTCAAAAGAGTGTGGGAGGCGAGCTGGGAGCAGATGT 423  
 QY 1898 GCTGCACTGGGAGGAGGCGCTCCACCTCTACTCTGCGAGCTGGAGCCAGTGC 1957  
 Db 424 GCTGCACTGGGAGGAGGCGCTCCACCTCTACTCTGCGAGCTGGAGCCAGTGC 483  
 QY 1958 CTGCTACGCTTTCACGAGCAGCTGGGCGCTTTGGCTGGTGGAGAGCCCTCAGCGT 2017  
 Db 484 CTGCTACGCTTTCACGAGCAGCTGGGCGCTTTGGCTGGTGGAGAGCCCTCAGCGT 543  
 QY 2018 GGTGCGGCCAAGGCTCAAGCTGCTTCTGTTTGGCCCGTGGCTGCACTCCCTCGA 2077  
 Db 544 GGTGCGGCCAAGGCTCAAGCTGCTTCTGTTTGGCCCGTGGCTGCACTCCCTCGA 603  
 QY 2078 GTACACATCCGGTCTACTGCTGCTGATGACACCCAGTGCCTCAAGGAGGTGTGCA 2137  
 Db 604 GTACACATCCGGTCTACTGCTGCTGATGACACCCAGTGCCTCAAGGAGGTGTGCA 663  
 QY 2138 GCTGGAGAAGCAGCTGGGAGGAGCAGCTGATCCAGGAGCCAGGCTCTGCACTCAAGGA 2197  
 Db 664 GCTGGAGAAGCAGCTGGGAGGAGCAGCTGATCCAGGAGCCAGGCTCTGCACTCAAGGA 723  
 QY 2198 CAGTTACCAACCTGGGCTATCCATCCAGTGTGCCAGCTCCCTGTGGAGGTAA 2257  
 Db 724 CAGTTACCAACCTGGGCTATCCATCCAGTGTGCCAGCTCCCTGTGGAGGTAA 783  
 QY 2258 GCTCCTTGTGAGTACAGGAGATCCCTTTATCATCTGGAATCGGACGCGAGCGTA 2317  
 Db 784 GCTCCTTGTGAGTACAGGAGATCCCTTTATCATCTGGAATCGGACGCGAGCGTA 843  
 QY 2318 CTTGCACTGACCTTACCTTGGAGCGTGTGAGCCCGAGCACTAGTACCTGGCTGCA 2377  
 Db 844 CTTGCACTGACCTTACCTTGGAGCGTGTGAGCCCGAGCACTAGTACCTGGCTGCA 903  
 QY 2378 GCTGTGGTGTGGAGGTGGAGGCGGAGCGGAGCTTCACTCACTCAACATCAC 2437  
 Db 904 GCTGTGGTGTGGAGGTGGAGGCGGAGCGGAGCTTCACTCACTCAACATCAC 963  
 QY 2438 CAAGGACACAAGTGTGCTGAGCTGCTGAGAGTGAAGCGGGGTCCAGCCCT 2497  
 Db 964 CAAGGACACAAGTGTGCTGAGCTGCTGAGAGTGAAGCGGGGTCCAGCCCT 1023  
 QY 2498 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGAGAGATATTTCCAGCCT 2557  
 Db 1024 GGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGAGAGATATTTCCAGCCT 1083  
 QY 2558 GGACCCACCTGTAGGGGGTGGAGTGGGAGCTTGGCCAGAACTCCACCTGGA 2617  
 Db 1084 GGACCCACCTGTAGGGGGTGGAGTGGGAGCTTGGCCAGAACTCCACCTGGA 1143  
 QY 2618 CAGGCATCTCAGCTTCTTGGCTTCAAGCCAGCCGCCACAGCCATGATCTTCACTGTG 2677  
 Db 1144 CAGGCATCTCAGCTTCTTGGCTTCAAGCCAGCCGCCACAGCCATGATCTTCACTGTG 1203  
 QY 2678 GGAGCGCGGCACTTCCCAACGGCACTCAGCCAGTGTGCTGAGAGTGGTGGACT 2737  
 Db 1204 GGAGCGCGGCACTTCCCAACGGCACTCAGCCAGTGTGCTGAGAGTGGTGGACT 1263  
 QY 2738 GGGCCAGCAGCAGCTGCGCTTTCACAGTGTGAGAGCTGAGTCTCAGGC 2789  
 Db 1264 GGGCCAGCAGCAGCTGCGCTTTCACAGTGTGAGAGCTGAGTCTCAGGC 1315

AAH99530  
 ID AAH99530 standard; cDNA; 1321 BP.  
 XX  
 AC AAH99530;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Human protein encoding cDNA sequence SEQ ID NO:365.  
 XX  
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 KW antibacterial; endocrine; cardiant; central nervous system; virucide;  
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
 KW antiaggregant; haemostatic; vulnery; antidiabetic; cyostatic;  
 KW dermatological; antiallergic; antiasthmatic; antidiabetic; cyostatic;  
 KW neuroprotective; antidepressant; neurotropic; antiparkinsonian; infection;  
 KW immunostimulant; gene therapy; antiseptic shock; pancreatitis;  
 KW antianaphylactic; rheumatoid arthritis; cardiac anaphylaxis; autoimmunity;  
 KW cardiac dysfunction; neuropathologic disorder; platelet disorder; asthma;  
 KW genetic disease; haematopoietic disorder; severe combined immunodeficiency;  
 KW thrombocytopaenia; osteoporosis; multiple sclerosis; depression;  
 KW allergic rhinitis; diabetes; Parkinson's disease; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 KW neurological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200153455-A2.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 22-DEC-2000; 2000WO-US035017.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-457603/49.  
 DR P-PSDB; AAM25589.  
 XX  
 PT Isolated human polynucleotides encoding polypeptides, useful for the  
 XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.  
 PS  
 CC Claim 1; Page 471; 1217pp; English.  
 CC  
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
 CC AAM25963. The proteins can have activities based on the tissues and cells  
 CC they are expressed in, such as: antinflammatory; antirheumatic;  
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;  
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;  
 CC antiulcer; osteopathic; dermatologic; antiallergic; antidiabetic;  
 CC antidiabetic; cyostatic; neuroprotective; antidepressant; neurotropic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antiseptic therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders  
 XX  
 SQ Sequence 1321 BP; 258 A; 440 C; 371 G; 252 T; 0 U; 0 Other;

Query Match				41.4%; Score 1193.6; DB 4; Length 1321;			
Best Local Similarity				97.8%; Pred. No. 1.4e-210;			
Matches 1295; Conservative				0; Mismatches 19; Indels 10; Gaps 8;			
QY	1476	CACACACTGACACACAGCTCTCCACACTCTGAGGCGGAGGAGTTCGTCTCCCGCTCTCC	1535				
DB	1	CACACACTGACACACAGCTCTCCACACTCTGAGGCGGAGGAGTTCGTCTCCCGCTCTCC	60				
QY	1536	ACCAGAACTACTTCCGCTCCCTCCCGGAGGACACAGACATGACTATGGAGACTTC	1595				
DB	61	ACCAGAACTACTTCCGCTCCCTCCCGGAGGACACAGACATGACTATGGAGACTTC	120				
QY	1596	AACCTCTCCGGGGCCGGCTGATGATCCCTAATACAGATATCAGCTCTCATCCCCCA	1655				
DB	121	AACCTCTCCGGGGCCGGCTGATGATCCCTAATACAGATATCAGCTCTCATCCCCCA	180				
QY	1656	GATGCCATACCCGAGGAGAGATCTATGAGATCTACCTCAGCTGCAAGCCGGAAGAC	1715				
DB	181	GATGCCATACCCGAGGAGAGATCTATGAGATCTACCTCAGCTGCAAGCCGGAAGAC	240				
QY	1716	GTGAGTTGGCCCTAGTGGCTGTACAGCCCTGTGAGTCCATCGTTAGCTGTGACCC	1775				
DB	241	GTGAGTTGGCCCTAGTGGCTGTACAGCCCTGTGAGTCCATCGTTAGCTGTGACCC	300				
QY	1776	CCT-GGGCTCTGCTACCCGGCCAGTATCCT-GGCTATGGACCACTGT-GGGGAGCCC	1832				
DB	301	CCTGGGCTCTGCTTACCCGGCCAGTATCCTGGGCTATGGACCACTGTGGGGAGCCC	360				
QY	1833	AGCCCTGACAGCT-GGAGCTCGCCCTCAAAAAGCAGTCTGTGAGGCGAGCTGGGAGCA	1891				
DB	361	AGCCCTGACAGCTGGGAGCTCGCCCTCAAAAAGCAGTCTGTGAGGCGAGCTGGG--A	417				
QY	1892	GGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACTCTACTACTGCCAGCTGGAGGC	1951				
DB	418	GGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACTCTACTACTGCCAGCTGGAGGC	477				
QY	1952	CAGTGCCTGCTAGTCTTCAACGAGCAGCTGGCGCGCTTTGCCCTGTGGGAGAGGCCCT	2011				
DB	478	CAGTGCCTGCTAGTCTTCAACGAGCAGCTGAGCGCGCTATGCCCTGGTGGGAGAGGCCCT	537				
QY	2012	CAGCGTGGCTGGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCCCTGCACCTC	2071				
DB	538	CAGCGTGGCTGGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCCCTGCACCTC	597				
QY	2072	CCTCGAGTCAACATCCGGGTCTACTGCTGTGATGACACCCAGTGCATCAGAGAGT	2131				
DB	598	CCTCGAGTCAACATCTACTGCTGTGATGACACCCAGTGCATCAGAGAGT	657				
QY	2132	GGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTCGCACTT	2191				
DB	658	GGTGAGCTGGAGAGCAGCTGCAGGAGCAGCTGATCCAGGAGCCACTGTACTGCACTT	717				
QY	2192	CAAGGACAGTACCAACCTGGCTATCCATCCAGATGTGCCAGTCTCCCTGTGGAA	2251				
DB	718	CAAGGACAGTACCAACCTGGCTATCCATCCAGATGTGCCAGTCTCCCTGTGGAA	777				
QY	2252	GAGTAAGCTCTTGTGAGTACAGAGATCCCTTTTATCATCTGGAATGGACGCA	2311				
DB	778	GAGTAAGCTCTTGTGAGTACAGAGATCCCTTTTATCATCTGGAATGGACGCA	837				
QY	2312	GGGTACTTGCACTGACCTTCAACCTGGAGCGTGTACGCCCCAGCACTAGTACCTGCG	2371				
DB	838	GGGTACTTGCACTGACCTTCAACCTGGAGCGTGTACGCCCCAGCACTAGTACCTGCG	897				
QY	2372	CTGCAAGCTGTGGGTGTGCGAGTGGAGGCGAGCGGAGAGCTTCAGATCAACTTCAA	2431				
DB	898	CTGCAAGCTGTGGGTGTGCGAGTGGAGGCGAGCGGAGAGCTTCAGATCAACTTCAA	957				
QY	2432	CATCACCAGGACACAAGTTTGTGAGCTGTGCTCTGGAGTGAAGCGGGGTGCC	2491				
DB	958	CATCACCAGGACACAAGTTTGTGAGCTGTGCTCTGGAGTGAAGCGGGGTGCC	1017				
QY	2492	AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATAATTTC	2551				
DB							
DB	1018	AGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATAATTTC	1077				
QY	2552	CAGCTTGAGACCCACCTCTGTAGGCGGGGTGCCAGTCTGGCGGACTCTGGCCAGAACTCCA	2611				
DB	1078	CAGCTTGAGACCCACCTCTGTAGGCGGGGTGCCAGTCTGGCGGACTCTGGCCAGAACTCCA	1137				
QY	2612	CCTGGAAGAGCAATCTCAGCTTCTTTGCTCTCAAGCCAGCCCCACAGCCATGATCTCAA	2671				
DB	1138	CCTGGAAGAGCAATCTCAGCTTCTTTGCTCTCAAGCCAGCCCCACAGCCATGATCTCAA	1197				
QY	2672	CCTGTGGGAGGCGGCGCACTTCCCAAGGCAACTCAGCAGCTGGTGCAGAGTGGC	2731				
DB	1198	CCTGTGGGAGGCGGCGCACTTCCCAAGGCAACTCAGCAGCTGGTGCAGAGTGGC	1257				
QY	2732	T-GGACTGGGCCAGCAGAGCGCTGGCCCTC-TTCACAGTG-TCCGAGGCTGAGTCTGAGG	2788				
DB	1258	TGGGACTGGGCCAGCAGAGCGGTGGCTTCTTTCACAGTGTTCCGAGGCTGAGTCTGAGG	1317				
QY	2789	CCGG 2792					
DB	1318	CCGG 1321					
RESULT 11							
AAS75738							
ID	AAS75738 standard; cDNA; 3646 BP.						
AC	AAS75738;						
XX							
DT	13-FEB-2002 (first entry)						
XX							
DE	DNA encoding novel human diagnostic protein #11542.						
XX							
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;						
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.						
XX							
OS	Homo sapiens.						
XX							
PN	W0200175067-A2.						
XX							
PD	11-OCT-2001.						
XX							
PF	30-MAR-2001; 2001WO-US008631.						
XX							
PR	31-MAR-2000; 2000US-00540217.						
PR	23-AUG-2000; 2000US-00649167.						
XX							
PA	(HYSE-) HYSEQ INC.						
XX							
PI	Drmanac RT, Liu C, Tang YT;						
XX							
DR	WPI; 2001-639362/73.						
DR	P-PSDB; ABG11551.						
XX							
PT	New isolated polynucleotide and encoded polypeptides, useful in						
PT	diagnostics, forensics, gene mapping, identification of mutations						
PT	responsible for genetic disorders or other traits and to assess						
PT	biodiversity.						
XX							
PS	Claim 1; SEQ ID NO 11542; 103pp; English.						
XX							
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)						
CC	sequences. (I) is useful as hybridisation probes, polymerase chain						
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,						
CC	and in recombinant production of (II). The polynucleotides are also used						
CC	in diagnostics as expressed sequence tags for identifying expressed						
CC	genes. (I) is useful in gene therapy techniques to restore normal						
CC	activity of (II) or to treat disease states involving (II). (II) is						
CC	useful for generating antibodies against it, detecting or quantitating a						
CC	polypeptide in tissue, as molecular weight markers and as a food						
CC	supplement. (II) and its binding partners are useful in medical imaging						
CC	of sites expressing (II). (I) and (II) are useful for treating disorders						

CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3646 BP; 930 A; 917 C; 921 G; 877 T; 0 U; 1 Other;

Query Match 31.8%; Score 916; DB 5; Length 3646;  
 Best Local Similarity 60.3%; Pred. No. 2.1e-159;  
 Matches 1657; Conservative 1; Mismatches 1016; Indels 72; Gaps 6;

QY 98 GCCCGGCTGTGGCCAGCGCTCTGGGCATAGTCCTCGCGCTTGGCTCGCGGCTCGGG 157  
 DB 220 GCTCGTGTACCTGCGCTGCGCTCTGCTAGCGGCGGCACTGGCTCCGCCCAAGA 279

QY 158 TGCCGACGAGTGCACCCGCTGGCCAAACCCAGTGCCTGTGCGCAACCCGACCTGCTTC 217  
 DB 280 TGATGACTTTTTCATGAATCCCAAGAACTTTTCTCTGTGATCCACCTGAGCCTGCG 339

QY 218 CCATTCCTGCTGGTGGAGCCCGAGAGTGTATCATGCTCAAGAAAGCAAGCCAGTGTCTGT 277  
 DB 340 ACATTTCTTATTGAGCTGAAGAAGCTTATTGTGAAGAAAGCAAGCTGTGAACCTGTA 399

QY 278 GTCCAGGCGGTGCGCCGACGAGATCTTCAAGTGCACCGGGAGTGGGTGCGCCA 337  
 DB 400 CTGTAAAGCAAGCCCTGCCACCCAGATCTATTCCAGTGTATAGTGAATGGTTATCA 459

QY 338 GGTGGACCACTGATGAGCGGACGACAGCGGAGCAGTGGTGAGCCGACCAATGGAGGT 397  
 DB 460 GAAGGACCAATAGTAGAAGAGTAGATGAATCTTCGGTCTCATTTGCGGGAAGT 519

QY 398 CCGCATTAATGTCTCAAGGACGAGTGTGAGAGGTGTGGGCTGAGGAACTACTGGTG 457  
 DB 520 GAGCATTTAGATTTTCGGCCAGCAAGTGAAGAACTCTTTGGACCTGAAGATTAAGTGTG 579

QY 458 CCAGTGTGTCATGAGTGTCTCGGACACCAAGAGTGCAGAGGCTCATATCCGAT 517  
 DB 580 CCAGTGTGTCATGAGTGTCTCGGACACCAAGAGTGCAGAGGCTCATATCCGAT 639

QY 518 AGCCAGATTTGGCAAGAACTTCGAGCAGGAGCCGCTGCGCAAGAGGTGTCCCTGGAGCA 577  
 DB 640 TGCAATATCTACGGAAGACATTTGAGCAGGAAACCCCTAGGAAGAGTGTCTTTGGAACA 699

QY 578 GGGCATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637  
 DB 700 GGAAGTCTTACTCAGTGTGCAACCACTGGAAGGATCCCACTGCTGAGGTGGAATGGTT 759

QY 638 CCGGAAACGAGACCTGTGTGAGCCGCTGCTGCAACCCCAATGATATACATCAACCGGAGCA 697  
 DB 760 GAAATATGAAGCAATATGATCCGTTGAAGATCGGAATTTTATATTAATGATCA 819

QY 698 CAGCTGTGTGTGCGACAGGCGCGCTTGTGTGACACGGCCAACTACCTGCTGGCGCAA 757  
 DB 820 CAACCTCATATAAAGCAGGCGCGCTTGTGTGATGATGCAATTAACCTGTGTGCGCAA 879

QY 758 GAAATGTGTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 817  
 DB 880 AAAATTTGTCGAGGAAAGTATCAACTGCGCACTGTCTAGTGTATGATCAACGGTGG 939

QY 818 GTGTGTCAGTGTGAGCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877  
 DB 940 CTGTGTCACCTGAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 999

QY 878 ACAGGACCGGAGCTGCAACCAACCGGCGCTTCTCAACCGGCGGCGCTTCTGTGAGGGCA 937  
 DB 1000 ACGTAAAGGACTTGTGACCAACCGGCGCACTCAATGAGGGGTGCTTCTGTGAGGGCA 1059

QY 938 GAATGTCCATGACCGCACCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 997  
 DB 1060 GAGTGTGAGAA---AATAGCCTGTACTACGTTATGCCAGTGAATGGAGTGGACGCC 1116

QY 998 GTGGAGCAAGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1057  
 DB 1117 ATGGAGCAAGTGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1176

QY 1058 TGACCAGCACCCCGCAACGAGGGGAGGAGTGCAGGGGCACTGACTTGACACCCCGCAA 1117  
 DB 1177 GCGCCAGCCCCCAAGAAATGGAGGCAAGGAGTGCAGGGCTCTCTTTCGAATCCAAGAA 1236

QY 1118 CTGTACCACTGACCTCTGTGTACACAGTGTCTTTCGCCCTGAGGACGTGGCCCTCTATGT 1177  
 DB 1237 CTGCACTGATGGCTTTTGCATGCAGACTGCTCTGATTCAGATGATGTTGCTCTCTATGT 1296

QY 1178 GGG---CCTCATCGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234  
 DB 1297 TGGGATTGTATAGCACTGATCGTTTGGCTGGCGATCTCTGTAGTGTGTGGCTTGTGTGT 1356

QY 1235 TTATTCGCCGAGAGAGGAGGCTGGACTCAGATGTGGCTGACTCGTCCATTTCTCACTC 1294  
 DB 1357 GTATCGGAAGATCATGCTGACTTTGAGTCAGATATTTATGACTCTTTCGCACTCATGG 1416

QY 1295 AGCTTTCAGCCCGTCAAGCATCAAGCCAGCAAGAGCAGCAACCCCACTGTCTCACT 1354  
 DB 1417 GGGCTTTTCACCTGTGACATCAAG-----GCAGCAAGCAAGATCTGTGTGGCTGT 1467

QY 1355 CCAGCGGAGCTCAGCACCAACCACTTACAGGGGAGTCTGTGCTCCCGGAGGATGG 1414  
 DB 1468 ACCCCAGAGCTCAGCTCAGTCAGCATGTACAGAGGACCTGTCTATGCCCTGCAATGA 1527

QY 1415 GCCAGCCCCAAGTTCAGCTCACCAGTGGGACCTGTCTCAGCCCCCTGGGTGGGCGCG 1474  
 DB 1528 CGTCTCAGCAAAATCCCAATGACCACTCTCCAACTCTGGATCCACTGCCCACTGAA 1587

QY 1475 CCACACACTCCACACA-----GCTCTCCACCTCTGAGCGGAGGAGTGTGT 1522  
 DB 1588 AATCAAGTGTACACACCTCAGTGTGTCTCCCCCAAGATGACCTCTCTGAGTTTAC 1647

QY 1523 CTCCCGCTCTCCACCCAGAACTATCTTCGCTCTCCCGGAGGACCCAGCACTGA- 1581  
 DB 1648 GTCCAGCTGTCCCTCAGATGACCCCAAGTGTGTGGAGATGAAGCCCTCAGCCCTGAA 1707

QY 1582 -----CCTATGGGACCTTCAACTT 1600

DB 1708 GAACCAAGTCTAGCAAGGAGACTGATCCATCTGTACCGCATTTGGAGCTTCACTC 1767

QY 1601 CCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGGCTCTCTATCCCCCAGATGC 1660

DB 1768 GCTGGAGGTCACTTATTTTCCCAATTCAGGAGTCACTTGTGTGATTTCCCGTGGGCG 1827

QY 1661 CATACCCCGAGGAGATCTATGATCTACCTCAGCTGCACAAAGCGGAGAGCTGAG 1720

DB 1828 CATTCGCAAGGAGAGTCTACGAATGTATGTGACTGTACACAGGAAAGAACTATGAG 1887

QY 1721 GTTCCCTAGCTGCTGTGAGACCTGCTGAGTCCCATCTGTAGCTGTGACCCCTG 1780

DB 1888 GCAACCAATGATGATCTCAGACACTTTTGACCCCTGTGTGAGCTGTGGCCCCCAGG 1947

QY 1781 CGTCTCTCAACCGGCGAGTCACTCTGGCTATGACCACTGTGGGAGGCCAGCCCTGA 1840

DB 1948 AGCTCTGTCAACCGGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2007

QY 1841 CAGCTGGAGCTGCGCTCAAAAGAGTGTGCGAGGAGCTGTGGAGAGGATGTCT 1900

DB 2008 GGAATGGAATACTGTCTCAAGAACAGGAGCACTGAGGAGAGTGGG---AGGATGTGT 2064

QY 1901 GCACCTGGCAGGAGGCGCTCTCCACTCTACTCTACTCTGCGAGCTGGAGGCGAGTGCCTG 1960

DB 2065 GGTGTGGGAGGAGAACTTCAACCCCTCTGTACATTAAGCTGGATGAGAGGCGCTG 2124

QY 1961 CTACGTCTTCAACCGAGAGCTGGGCGGCTTTGCGCTGTGGAGAGGCGCTCAGCGTGC 2020

Db 2125 CCACATCTCTCAGAGAACCTCAGCACCTACGCTCTGTTAGGACATTCACACCAAGC 2184  
Qy 2021 TGCCGCCAAGCCCTCAAGCTCTTCTGTTTCGGCGGTTGCTGCTGCTCTCTCGAGTA 2080  
Db 2185 GGCTGCAAGCCCTCAAGCTGCTCTTGGGCCCCCTGCTGCTCTCTCTGCTGAGTA 2244  
Qy 2081 CAAATCCGGTCTACTGCTGCTCATGACACCCAGATGACCTCAAGGAGGTGGTGCAGCT 2140  
Db 2245 CAGCATCCGAGTCTACTGCTGATGACACCCAGATGCCCTGAAGGAATTTTACATCT 2304  
Qy 2141 GGAGAGAGCTGGGGGACAGCTGATCCAGAGACCGGGTCTGCTCAAGGACAG 2200  
Db 2305 TGAGAGACAGCGGAGGACAGCTCTTAGAAGAACCTAAGGCTCTTCAATTTAAGGCGAG 2364  
Qy 2201 TTACCACAACTTCCCTCTATCCATCCAGATGTCGCCAGCTCCCTGTGGAAGAGTAACT 2260  
Db 2365 CACCCACAACTTCCCTCTGCTCAATTCAGATATGCCCATTCCTCTGGAAGAGCAAT 2424  
Qy 2261 CTTGTGAGTACAGGAGATCCCTTTTATCATCTGGAATGGACGCGGTACTT 2320  
Db 2425 GCTGGCTAAATATCAGGAATTCATTTTACCATGTTTGGAGTGGATCTCAAGGAACCT 2484  
Qy 2321 GCAGTGCACCTTCACTCTGGAGCTGTGACGCCAGCTAGTACCTGCTGCTGCAAGCT 2380  
Db 2485 GCAGTGCACCTTCACTCTGGAGAGATTAGCTCTGAACACAGTGGAGCTGTTGCAACT 2544  
Qy 2381 GTGGGTGTGGCAGGTGGAGGCGGACGGCAGAGCTTCAGCATCACTCAACATCACCA 2440  
Db 2545 CTGTGTGGCAGGTGGAGGAGAGGCGAGATCTTCAGCTCACTGACCGGTGTGAGA 2604  
Qy 2441 GGACAAAGGTTTCTGAGCTGTGCTGTGGAGAGTGAAGCGGGGTCCAGCGCTGGT 2500  
Db 2605 GGAACCTACTGGCATCATTTGCGCTGCTGATCTCTGCGAACCATCACACGGTTCAC 2664  
Qy 2501 GGGCCCCAGGCTTCAAGATCCCTTCTCTCATTCGGCAGAGATAATTTCCAGCTGGA 2560  
Db 2665 GGGGCCAGTGTCTTACAGATCCCTTCTCTCATTCGGCAGAGCTGTGAGCAGCTTGA 2724  
Qy 2561 CCCACCTGTAGCGGGGTCGACTGGGAGCTCTGSCCCAGAACTCCACCTGGAAGCT 2620  
Db 2725 TGCCCCCAGACAGAGGCCATGACTGGAGATGCTGGCCCCAATGCTGAACCTGGACAG 2784  
Qy 2621 CACTCAGCTCTTTGCTCCCAAGCCAGCCACAGCCATGATCTCAACTGTGGGA 2680  
Db 2785 GTACTTGAATTAATTTGGCCACCAATCCAGCCCAACTGGGTAACTCTGATCTTTGGGA 2844  
Qy 2681 GCGCGGCACTTCCCAAGGCACTCAGCAGCTGCTGCGAGCTGCTGCGAGCTGGAGTGG 2740  
Db 2845 AGCAGAGAACTTCCAGATGGAACCTGAGCATGCTGGCAGCTGTCTTGAAGAAATGGG 2904  
Qy 2741 CCAGCCAGAGCTGGCTCTTTCAGTGTGCGAGGCTGAGTGTGA 2786  
Db 2905 AAGACATGAACGGTGTCTCTTACAGAGAGGAGGAGTATTA 2950

RESULT 12

ABT06279

ID ABT06279 standard; cDNA; 2860 BP.

XX

AC ABT06279;

XX

DT 24-OCT-2002 (first entry)

XX

DE Human NOV1a coding sequence.

XX

KW Human; NOVX; autoimmune disease; cancer; infection; inflammatory disease;  
KW storage disorder; muscle disorder; neurodegenerative disorder; nontropic;  
KW developmental defect; neuroprotective; antiparkinsonian; hypotensive;  
KW hypertensive; haemostatic; cardiant; antianigral; dermatological;  
KW immunosuppressive; antiinflammatory; virucide; antibacterial; anti-HIV;  
KW antiparasitic; antiallergic; antiasthmatic; antirheumatic; antiarthritic;  
KW vulnary; anorectic; antidiabetic; immunomodulator; antipsoriatic;

KW nephrotropic; kerolytic; antiulcer; cerebroprotective; anticonvulsant;  
KW antifertility; antimanic; antidepressant; metabolic; cytostatic;  
XX tranquilizer; analgesic; gene; ss.  
OS Homo sapiens.  
XX WO200257450-A2.  
XX 25-JUL-2002.  
XX 29-NOV-2001; 2001WO-US048922.  
XX 29-NOV-2000; 2000US-0253834P.  
PR 30-NOV-2000; 2000US-0250926P.  
PR 25-JAN-2001; 2001US-0264180P.  
PR 20-AUG-2001; 2001US-0313656P.  
PR 05-OCT-2001; 2001US-0327456P.  
PR 28-NOV-2001; 2001US-00327456.  
XX (CURA-) CURAGEN CORP.

PA Edinger S, Macdougall JR, Millet I, Ellerman K, Stone DJ;

PI Gerlach V, Grosse WM, Alsobrook JP, Lepley DM, Rieger D, Burgess CE;  
PI Casman SJ, Spytek KA, Boldog FI, Li L, Padigaru M, Mishra V;  
PI Patturajan M, Shenoy S, Rastelli L, Tchervnev VT, Vernet CM;  
PI Zehusen BD, Malyankar UM, Guo X, Miller CE, Gangolli EA;

DR WPI; 2002-590741/63.

XX P-PSDB; MAO18734.

PT Novel isolated polypeptide, designated NOVX, useful for treating or  
PT preventing in NOVX-associated disorders e.g. cardiomyopathy,  
PT atherosclerosis, diabetes, cancer, allergy, asthma, Crohn's disease.

PS Claim 9; Page 11; 353pp; English.

XX The present invention provides the protein and coding sequences of  
CC several novel human proteins, designated NOVX. These can be used in the  
CC treatment of, amongst others, cancers, autoimmune diseases, infections,  
CC inflammatory diseases, storage disorders, muscle disorders,  
CC neurodegenerative diseases and developmental defects. The present  
CC sequence is a coding sequence of the invention

SQ Sequence 2860 BP; 555 A; 950 C; 860 G; 495 T; 0 U; 0 Other;

Query Match 30.9%; Score 889; DB 6; Length 2860;

Best Local Similarity 60.7%; Pred. No. 1.9e-154; Mismatches 1637; Conservative 0; Matches 960; Indels 102; Gaps 7;

Qy 184 ACCAGTGGCTGTGTCACACCCGGACCTGCTTCCCCACTTCTCTGGTGGAGCCGAGGATG 243

Db 168 ACTCTTCCCGTCAGCGCCAGCAGAGCCGCTGCGCTACTTCTTGCAGGACACAGGCG 227

Qy 244 TGTACATCTCAAGACACAGCCAGTGTCTTGTGTGAAGCCGCTGCCCGCCACGAGA 303

Db 228 CTTACATTTGAAGAACAAAGCCTGTGGAGCTCTCGCTGCCGCTTCCCGCCACACAGA 287

Qy 304 TCTTCTTCAAGTGCAACGGGGAGTGGTGGCGAGTGGAGTCCGCTTAAATGTCTCAAGGACAGCAG 363

Db 288 TCTACTTCAAGTGCAACGGCGAGTGGTGGCGAGTGGAGTGGAGTGGAGTGGAGTGGAG 347

Qy 364 CAGAGGGAGAGTGGTGGAGCCAGCATGGAGTCCGCTTAAATGTCTCAAGGACAGCAG 423

Db 348 TGGATGAGCCACCGCTTCTGCGGTGCGAGGTGCAGATCGAGTGTGCGGCGAGCAGG 407

Qy 424 TCGAAGAGTGTTCGGGCTGGAGGATTAATGTGTGCGAGTGTGCGAGTGTGCGAGTGTGCG 483

Db 408 TGGAGAGTCTTTGGGCTGGAGGATTAATGTGTGCGAGTGTGCGAGTGTGCGAGTGTGCG 467

Qy 484 GCACCAACAGAGTGCAGAGGCTTACATCCGATAGCCAGATTGGCAGAGTGTGAGC 543

Db 468 GCACCAACAGAGTGCAGAGGCTTACATCCGATAGCCAGTGTGCGCTTACGCGAAGAACTTCGATC 527







1185 TGGCAATCTCATGCGGCTGGGGTGGTGTACCGCCGCACTGCGGTGACTTCGACA 1244  
 1264 CAGATGGGTGACTGCTCCATT---CTCAGCTCAGGCTTCCAGCCCGTCCAGCATCAAGC 1320  
 1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGTTCACACCCGCTCAACTTAAAG 1304  
 1321 CCAGCAAGAGACAAACCCGCTCTGCT-----CAGCATCAGCGGACCTCAGCACCA 1374  
 1305 CGGCAAGGCCAGTAACCGGAGCTCTACACCCCTCTGTGCTCTGACCTCAGACCCA 1364  
 1375 CCACCACTTACAGGAGCTCTCTGCTCCCGGAGGATGGCCAGCCCAAGTTCAGC 1434  
 1365 GCGCCGCGCATCTACCGGAGACCTGTATGCTCCCTGAGGACTCAACGACAAAATCCCCA 1424  
 1435 TCACCAATGGGCACTGCTCAGCCCC-----1460  
 1425 TGACCACTCTCTCTGCTGAGACCCCTTACCAGCCTTAAAGTCAAGGTCTACAGTCCA 1484  
 1461 -----CTGGGTGGCGCGCC 1476  
 1485 GCACCAAGGCTCTGGGCCAGGCTTGACAGATGGGGTGCACCTCTGCGGCTTTGCGCG 1544  
 1477 ACACACTGCACACAGCTCTCCACCTCTGAGGCGGAGGAGTTCGTCTCCCGCTCTTCCA 1536  
 1545 CTGSCACATACCTTAGGATTTGCGCCGGGACACCCACTCTCTGCACTGCGGAGCGCCA 1604  
 1537 CCCAGAACTAC-----TTCCGCTCTCCCTGCGCGGAGCACAGCAACATCACTATCGGA 1590  
 1605 GCCTCGGTTCCAGCAGCTCTTGGGCTGCGCCGAGACCCAGGAGCAGCGTCAAGCGCA 1664  
 1591 CTTTCAACTCTCTCGGGGCGGCTGATGATCCCTTAATACAGGTATCAGCTCTCTATCC 1650  
 1665 CTTTGGCTGCTGGGTGGAGGCTCAGCATCCCGCCGACAGGGGTGAGTTCGTGCTGTC 1724  
 1651 CCCAGATGCATACCCGAGGAGAGATCTATGATCTACTCACGCTGCACAAAGCGG 1710  
 1725 CCAATGAGCCATTCCTCCAGGCGAAGTTCTACGAGATGTATCTACTCATCAACAGGAG 1784  
 1711 AAGAGTGAGGTGCCCTAGCTGGCTGTAGACCCCTGAGTCCCATGCTTACGTGTG 1770  
 1785 AAAGTACCTGCGCTTTTCAAGGGGACCCAGACAGTATTGAGCCCTCGGTGACCTGTG 1844  
 1771 GACCCCTGGGCTCTGCTCACCCTGAGTCTATCTGCTGCTGATGACACCTGTCGGGAGC 1830  
 1845 GACCCACAGGCTCTGCTGTGCGCCGCGCTCATCTCCATCACCATGCCCTGTCGCCAAG 1904  
 1831 CCAGCCTGACAGCTGAGCTGCGCTCAAAAAGCAGTCTGTGGAGGCGAGCTGGGAGC 1890  
 1905 TCAGTCCCGTACGTGATCTTTTTCAGTCAAGACCCAGGCCACAGGCGCTGGGAG- 1963  
 1891 AGGATGTCTGACCTGCGGAGGAGGCGGCTCCACCTCTACTACTGCGCAGCTGGAGG 1950  
 1964 --GAGGTGTGACCTGATGAGGAGACCTGAAACACACCTGCTACTGCCAGCTGGAGC 2021  
 1951 CAGTGTCTGCTACTCTTACAGGAGCTGCGGCGCTTTGCGCTGCTGGAGGAGGCGCC 2010  
 2022 CCAGGCGCTGTCACTCTGCTGACAGCTGGGCACTTACGTGTTCAGGCGGAGTCTCT 2081  
 2011 TCAGCTGGCTGCGGCCAGGCGCTCAAGCTGCTTCTGTTTGGCGGCTGGCTGCACT 2070  
 2082 ATTCCGCTGAGCAGTCAAGCGGCTCAGCTGGCGCTTTCGCCCGCCCTCTGCACCT 2141  
 2071 CCCTGAGTCAACATCCGGGTCTACTGCTGCTGATGACACCCCATGATCAAGGAGG 2130  
 2142 CCCTGAGTACAGCTCCGGGTCTACTGCTGGAGGACACGCTGTAGCAGTGAAGAGG 2201  
 2131 TGGTCACTGAGAGAGCTGGGGGAGCAGCTGATCCAGGAGCACCGGCTCTGCACT 2190  
 2202 TGCTGAGCTGAGCGGAGCTCTGGGCGGATATCTTGGTGGAGGAGCGGAAACCGCTAATGT 2261  
 2191 TCAAGGACAGTGTACACACCTGCGCTTATCCATCCAGATGTGCCAGCTCCCTGTGGA 2250

2262 TCAAGGACAGTTACCAAACTTGGCGCTCTCCCTCCATGACCTCCCCCATGCCCATTTGA 2321  
 2251 AGAGTAAGCTCTGTGCTAGCTACAGAGATCCCTTTTATCACATCTGGAATGGCAGCG 2310  
 2322 GGAGCAAGCTGTGGCCAAATACAGAGATCCCTTCTATCACATTTGGAGTGGCAGCC 2381  
 2311 AGCGGTACTTGCACCTTCCACCTTGAGCGTGTAGCCCCAGCACTAGTGAGCTGG 2370  
 2382 AGAAGGCCCTCCATGCTCACTTTTCCCTTGAGAGGACAGCTTGGCTCCACAGAGTCA 2441  
 2371 CTTGCAAGCTGTGGGTGTGGAGTGGAGGCGACGAGCAGAGCTTCAGCATCACTTCA 2430  
 2442 CTTGCAAGATCTGCTGCGGCAAGTGAAGGAGGAGGCGGAGGATATTCAGCTGCATACCA 2501  
 2431 ACATCAACCAAG---GACACAAGGTTTGTGAGCTGTGCTGAGAGTGAAGCGGGG 2487  
 2502 CTTGCGCAGAGACACCTGTGCTGCTGGAGACTCTCTGCTCTGCTGCGCAGCACTG 2561  
 2488 TCCAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTTCCTCATTTGGCAGAGATAA 2547  
 2562 TCACCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCGCAGAGATAT 2621  
 2548 TTTCCAGCTTGGAGCCACCTGTAGGGGGTCCGACTGGCGGACTCTGCGCCAGAAAC 2607  
 2622 GCAACAGCTTAGATGCCCCCACTCAAGCGGCAATGACTGGCGGATGTTAGCAGAGAGC 2681  
 2608 TCCACCTGGAGACGCACTCTCAGCTTTTTCCTTCCAAAGCCAGCCCCACAGCATGATCC 2667  
 2682 TCTCTATGAGCCGCTACCTGAATATTCTTCCACCAAGCGAGCCCGGCTGATCC 2741  
 2668 TCAACCTGTGGAGGCGCGCACTTCCCAACCGCAACCTCAGCCAGCTGGCTGCAGAG 2727  
 2742 TGGACCTCTGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCTGGCGAGTGCCT 2801  
 2728 TGGCTGAGCTGGGCGCAGCAGCGTCCCTTCTTCAAGTGTGGAGGCTGAGTGTGA 2786  
 2802 TGGAGGAGATGGGCAAGAGTGTGCTGGTGGTGTGGCCACCGAGCGGAGCTGCTGA 2860

RESULT 14

ABK92062  
 ID ABK92062 standard; DNA; 2995 BP.

AC ABK92062;

XX  
 DT 14-AUG-2002 (first entry)

XX DNA encoding novel UNC5 receptor-like protein.

DE Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;  
 XX behavioral disorder; valve disease; endocrine disorder; heart disorder;  
 KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW immune disorder; haematopoietic disorder; dyslipidaemia;  
 KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;  
 KW gene therapy; transgenic animal; human; gene; ds.

XX Homo sapiens.

OS WO200229058-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US031248.

XX 05-OCT-2000; 2000US-0238323P.

PR 06-OCT-2000; 2000US-0238325P.

PR 06-OCT-2000; 2000US-0238372P.

PR 06-OCT-2000; 2000US-0238373P.

PR 06-OCT-2000; 2000US-0238379P.

PR 06-OCT-2000; 2000US-0238382P.

PR 06-OCT-2000; 2000US-0238383P.

PR 06-OCT-2000; 2000US-0238384P.





Db 1401 AGCCAGCGCGGATCTACCGGAGACCCGGTGTATGCCCTGCAGGACTCCACCGACAAT 1460  
QY 1430 CCAGCTCACCAATGGGACCTCTAGCCCC----- 1460  
Db 1461 CCCATGACCAACTCTCTCTGCTGAGACCCCTTACCAGCCTTAAAGTCAAGGTCTACAG 1520  
QY 1461 -----CTGGGTGGCG 1471  
Db 1521 CTCACGACACCGGCTCTGGGCCAGCCCTGGCAGATGGGGCTGACCTGCTGGGGTCTT 1580  
QY 1472 CGCCACACACTGACACAGCTCTCCACCTCTGAGGGCGAGAGTTCGTCTCCGCGCT 1531  
Db 1581 GCGCCTGGCACAATCCCTAGCAATTTGCGCGGACACCCACTTCTGACCTGGGCGAG 1640  
QY 1532 CTCACCCC-----AGAACTACTTCCGCTCCCTGCCCCGAGGACACAGCAATGACCTA 1585  
Db 1641 CGCCAGCTCGGTTCCTCCAGCAGCTCTTGGGCTCTGCGCGAGACCCAGGAGAGCAGCGTCAG 1700  
QY 1586 TGGGACCTTCAACTTCTGCGGGGCGGCTGATGATCCCTTAATACAGGTATCAGCCTCCT 1645  
Db 1701 CGGCACTTTGGCTGCTGGGTGGGAGCTCAGCATCCCCGGGACAGGTGTAGCTTGCT 1760  
QY 1646 CATCCCCCAGAGCCATACCCGAGGAAAGATCTATGAGATCTACCTCAGCTGCACAA 1705  
Db 1761 GGTGCCAATGAGAGCCANTCCCGAGGCAAGTCTACGAGATGTATCTACTCATCACAA 1820  
QY 1706 GCGGAGAGAGCTGAGGTGGCTTCCCTAGCTGGCTGTACAGCCCTGCTGAGTCCCATCGTTAG 1765  
Db 1821 GGCAGAAAGTACCTCGCGCTTTTCAAGAGGACCCAGACAGTATTGAGCCCTCGGTGAC 1880  
QY 1766 CTGTGAGACCCCTGGGCTGCTGCTCACCAGGCGAGTATCTCTGCTATGACCACTGTGG 1825  
Db 1881 CTGTGAGACCAAGGCTCTGCTGTGCGCCCGCTCATCTCCACCATGCCCCACTGTGC 1940  
QY 1826 GGAGCCAGCCCTGACAGCTGGAGCTGGCGCTCAAAAAGCAGTGTGCGAGGCGAGCTG 1885  
Db 1941 CGAAGTCAGTCCGCTGACTGAGTCTTTTACGTCAAGCCAGGCCCCACAGGCCCACTG 2000  
QY 1886 GGAGCAGGATGTGTGCACCTGGGCGAGAGGCGCCCTCCACCTCTACTACTGCCAGCT 1945  
Db 2001 GGAGCAGGAGTGTGTGACCTGGATGAGGACCTGAAACACACCTCTACTGCCAGCT 2060  
QY 1946 GGAGCCAGTCTGTGCTGCTTACCGAGCAGCTGGGCGCTTCCCTCGTGGGAGA 2005  
Db 2061 GGAGCCAGGCGCTGTACATCTCTGCTGGACAGCTGGGACCTACGTTTCAAGGGCGA 2120  
QY 2066 GGCCCTCAGGTGCTGCGCAGAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCTG 2065  
Db 2121 GTCTATTCCGCTCAGCAGTCAAGCGCTCCAGCTGGCGCTTTCGCCCCCGCCCTCTG 2180  
QY 2066 CACCTCCCTCGAGTACAAATPCGGGTCTACTGCTGTGATGACACCCACGATGACTCAA 2125  
Db 2181 CACCTCCCTGAGTACAGCTCCGCGTCTACTGCTGGAGGACACGCTGTAGCACTGAA 2240  
QY 2126 GGAGGTGTGCTGAGTGGAGAGCAGCTGGGAGGACAGCTGATCCAGGAGCACAGGTCT 2185  
Db 2241 GGAGGTGTGAGTGGAGCGGATCTGCGCGGATATCTTGTGGAGGAGCGAAGCCCT 2300  
QY 2186 GCACCTCAAGACAGTTTACCAACACTCGGCTATCCATCCACGATGTGCCAGCTCCCT 2245  
Db 2301 AATGTTCAAGACAGTTTACCAACACTCGGCTCTCCCTCCATGACCTCCCCATGCCA 2360  
QY 2246 GTGGAAGATGAGTCTCTGTGAGTACAGAGATCCCTTTTATCACTTGAATGG 2305  
Db 2361 TTGGAGGAGCAAGTGTCTGGCCAAATACAGAGATCCCTTCTATCATATTTGAGTGG 2420  
QY 2306 CAGCAGCGGTACTTGTGACTGACCTTCCCTCGAGGTGTGAGCGGAGCAGCTAGTGA 2365  
Db 2421 CAGCCAGAGAGGCGCTTCCACTGACCTTTTCCCTTGAGAGGACAGCTGGGCTCCACAGA 2480  
QY 2366 CCTGCGCTCAAGCTGTGGGTGTGGAGGTGGAGGCGGACGCGGAGAGCTTCAGATCAA 2425  
Db 2481 GCTCACCTGCAGATCTGCTGCGGCAAGTGAAGGAGGAGGCGCAGATATTCAGTGA 2540

QY 2426 CTTCAACATCACCAAGGACA---CAAGGTTTGTGAGCTGCTGGCTCTCGAGAGTGAAGC 2482  
Db 2541 TACCACCTTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAG 2600  
QY 2483 GGGGGTCCAGAGCCCTGCTGGGGCCCCAGTCCCTTCAAGATCCCTTCTCATTTGGGAGAA 2542  
Db 2601 CACTGTCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAA 2660  
QY 2543 GATAATTTCCAGCTGGACCCACCTGTAGGGGGTGGCGACTGGCGGACTCTGGCCCA 2602  
Db 2661 GATATCAACAGCCTAGATGCCCCCAACTCAAGGGGCAATGACTGGCGGATGTAGCACA 2720  
QY 2603 GAAACTCCACCTTGCAGAGCCTCTCAGCTTCTTTTGGCTTCCAGGCCAGCCCCCAGCCAT 2662  
Db 2721 GAAGCTCTCTATGACCGGTACTGTAATTTTGGCCACCAAGCGAGCCCCACGGGTGT 2780  
QY 2663 GATCCTCAACCTGTGGAGCGCGGACATTTCCCAAGCGCACTTCAGCCAGCTGGCTGC 2722  
Db 2781 GATCCTGGACCTCTGGGAAGCTCTGAGCAGGACGATGGGGACCTCAACAGCCTGGCGAG 2840  
QY 2723 AGCAGTGGCTGGAAGTGGGCGAGCAGCAGCTGGCCTTTTCCAGAGTGTGCGAGGCTGAGTG 2782  
Db 2841 TGCTTTGGAGAGATGGCAAGAGTGAATGCTGGTGGCTGTGGCCACCGACGGGACTG 2900  
QY 2783 CTGAGGC 2789  
Db 2901 CTGAGCC 2907

RESULT 15  
ABQ93898  
ID ABQ93898 standard; DNA; 2895 BP.  
XX  
AC ABQ93898;  
XX  
DT 01-NOV-2002 (first entry)  
XX  
DE Human transmembrane receptor UNC5H2-like NOV11 DNA, SEQ ID NO:37.  
XX  
KW Human; NOVX; neurological disorder; Alzheimer's disease;  
KW Huntington's disease; Parkinson's disease; pain; behavioural disorder;  
KW addition; tuberosus sclerosis; cancer; immune disorder; allergy;  
KW autoimmune disease; myasthenia gravis; asthma; arthritis; diabetes;  
KW thyroiditis; cardiovascular disease; hypertension; reproductive disorder;  
KW endometriosis; incontinence; psoriasis; scleroderma; alopecia; ulcer;  
KW pancreatitis; cirrhosis; glomerular endotheliosis; bacterial infection;  
KW polycystic kidney disease; endocrine disorder; obesity; cardiomyopathy;  
KW atherosclerosis; cell signal processing-related disorder;  
KW metabolic pathway regulation disorder; cytostatic; neuroprotective;  
KW antiinflammatory; immunosuppressive; analgesic; antithrombotic;  
KW dermatological; antibacterial; antiarthritic; hepatotropic; neurogenesis;  
KW differentiation; proliferation; motility; haematopoiesis; wound healing;  
KW angiogenesis; forensic biology; transgenic animal; drug screening;  
KW gene therapy; NOV11; transmembrane receptor UNC5H2-like; chromosome 10;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
Key Location/Qualifiers  
5'UTR 1..30  
FT /\*tag= a  
FT 31..2868  
FT /\*tag= b  
FT /\*tag= "NOV11"  
FT /product= 31..108  
FT sig\_peptide  
FT /\*tag= c  
FT 109..2865  
FT mat\_peptide  
FT /\*tag= d  
FT /product= "Mature NOV11 protein"  
FT 2869..2895  
FT 3'UTR  
FT /\*tag= e  
XX

PN

WO200253742-A2.

XX

11-JUL-2002.

XX

07-JAN-2002; 2002WO-US000375.

XX

05-JAN-2001; 2001US-0260018P.

PR

08-JAN-2001; 2001US-0260360P.

PR

28-FEB-2001; 2001US-0272411P.

PR

02-MAR-2001; 2001US-0272817P.

PR

05-JUL-2001; 2001US-0303231P.

PR

12-JUL-2001; 2001US-0305060P.

PR

10-SEP-2001; 2001US-0318405P.

PR

12-SEP-2001; 2001US-0318700P.

PR

04-JAN-2002; 2002US-00037417.

XX

(CURA-) CURAGEN CORP.

PA

XX

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

PI

Kekuda R, Alsobrook JP, Tchernev VT, Liu X, Spytek KA; Patturajan M, Grosse WM, Lepley DM, Burgess CE, Vermet CAM, Li L; Gorman L, Edinger S, Sciore P, Ellerman K, Malyankar U; Rothenberg M, Stone D, Boldog F, Guo X, Shenoy S, Anderson D; Padigaru M, Taupier RJ, Miller CB, Eisen A; WPI; 2002-583619/62. P-PSDB; ABB09520.

Novel polypeptides and nucleic acids homologous to transmembrane receptor, thymosin, neuromodulin-like family of proteins for diagnosing, treating cancer, atherosclerosis, neurological, skin and autoimmune disorders.

Claim 9a; Page 121; 323pp; English.

The invention relates to 24 novel human proteins designated NOV1-NOV14 (ABB09501-ABB09524), collectively referred to as NOVX proteins, and nucleic acids encoding them (ABB093879-ABB093902). NOVX proteins and nucleotides are useful in the treatment, diagnosis or prevention of NOVX-associated disorders or in the manufacture of a medicament for treating such disorders, with specific applications described for each of the 24 NOVX proteins, based on their homology to known proteins. Various disorders are associated with NOVX proteins including neurological disorders (e.g., Alzheimer's, Huntington's and Parkinson's diseases), pain, behavioural disorders, addiction, tuberosclerosis, cancers (e.g., colorectal cancer, leukaemia and osteosarcoma), immune disorders (e.g., allergies and autoimmune diseases), myasthenia gravis, asthma, various forms of arthritis, diabetes, thyroiditis, cardiovascular disease (e.g., hypertension), reproductive disorders, endometriosis, incontinence, psoriasis, scleroderma, alopecia, ulcers, pancreatitis, cirrhosis, glomerular endotheliosis, polycystic kidney disease, endocrine disorders, obesity, bacterial infections and particularly cardiomyopathy, atherosclerosis, cell signal processing-related disorders and disorders of metabolic pathway regulation. NOVX nucleic acids and polypeptides may be used to identify cellular receptors or downstream effectors which bind to a NOVX protein, and are also useful as targets for the identification of small molecules that modulate or inhibit processes such as neurogenesis, cell differentiation, cell motility, cellular proliferation, haematopoiesis, wound healing and angiogenesis. NOVX nucleic acid sequences can be used to identify a cell or tissue type and are useful as a source of primers or probes for forensic biology and for identifying and cloning NOVX homologues in other cell types. Cells comprising NOVX nucleic acids are useful for producing non-human transgenic animals which are useful for studying the function and activity of NOVX proteins and for identifying and evaluating modulators of NOVX activity. The present sequence represents DNA encoding the transmembrane receptor UNC5H2-like protein NOV11. The gene encoding NOV11 is located on chromosome 10.

Sequence 2895 BP; 557 A; 960 C; 854 G; 524 T; 0 U; 0 Other;

Query Match 30.1%; Score 866.4; DB 6; Length 2895;

Best Local Similarity 60.6%; Pred. No. 2.8e-150;

Matches 1659; Conservative 0; Mismatches 941; Indels 138; Gaps 9;

QY 184 ACCAGTGCCTGGTGCCAAACCCGAGCCTGCTTCCCACTTCTGGTGAGCCGAGGATG 243  
DB 140 ACTCTTCCCGTCAGCGCCAGCAGAGCGCTGCTCTTCTTCTGAGGAGCCACAGGACG 199  
QY 244 TGTACATCGTCAAGAACAAAGCCAGTGTCTGTTGTGTGCAAGGCGGTGCGCCGACGACG 303  
DB 200 CTTACTTGTGAGAAACAAGCCTGTGAGCTCCGCTGCGCGCTTCCCGCCACACAGA 259  
QY 304 TCTTCTTCAAGTGCACAGCGGAGTGGGTGCGCCAGGTGGACACGTCATCGAGCGCAGCA 363  
DB 260 TCTACTTCAAGTGCACAGCGGAGTGGGTGCGCCAGGAGCCACGTCACACAGGAGGCC 319  
QY 364 CAGACGGAGCAGTGGTGAGCCGACCATGGAGTCCGATTTAATGTCTCAAGGAGCAGG 423  
DB 320 TGGATGAGGCCACCGGTCTCGGGTGCAGATCGAGGTGTGCGCGCAGGAGG 379  
QY 424 TCGAGAGGTGTTCCGGCTGGAGGAATACTGTGTCAGTGGTGGCATGAGGCTCTCTGG 483  
DB 380 TGGAGGAGCTCTTTGGGCTGGAGGATTAATGTGTCAGTGGTGGCTGGAGCTCCGCGG 439  
QY 484 GCACCAACCAAGAGTCAAGAGCCCTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGC 543  
DB 440 GCACCAACCAAGAGTCCCGAGCCTACGTCGCAATCGCTGTGCGCAAGAACTTCGATC 499  
QY 544 AGAGCGCTGGCCAAAGGAGGTGTCCTGGAGCAGGAGCATGTCGTGCTGCTGCGCTCCAC 603  
DB 500 AGGAGCTCTGGGCAAGGAGGTGCCCTGGACCATGAGGTTCTCTGTCAGTGGCGCCGC 559  
QY 604 CGAGAGGCAATCCCTCCAGCGGAGTGGGTCCGAAACGAGGACCTGGTGGAGCCGT 663  
DB 560 CGAGGAGGTGCTGTCGCGAGGTGGAAATGCGTCAAGATGAGGATGTTCATCGACCCCA 619  
QY 664 CCTTGGACCCCAATGTATACATCACGCGGAGCAGACGCTGGTGGCAGACGCGCGCC 723  
DB 620 CCCAGAGACCAACTTCTGCTCATCACCACACCAACCTCATCTCCGACGAGCGCGCC 679  
QY 724 TTGCTGACCGGCAACTACACTGCTGGCCCAAGAACATCGTGGCAGCTGCGCGCAGCG 783  
DB 680 TGTGCGACACTGCAACTATACCTGCTGGCCAAAGAACATCGTGGCAGGAGTGTCTAC 739  
QY 784 CCTCGCTGCTGTCATCGTCTAGCTGAAAGTGGTGGTGGAGTGGAGCGAGTGTGCG 843  
DB 740 CCACTGCCACCGTCACTGCTAGTGAATGGCGGCTGGTCCAGCTGGGAGAGTGTCTAC 799  
QY 844 TGTGACGCGCCAGCTGTGGCGCGGTGTCAGAAACGAGCGGAGCTGCACCAACCCCG 903  
DB 800 CCGTGTCCAAACCGCTGTGCGCGAGGTGCGCAGAGCGCACCGGACCTGCACCAACCCCG 859  
QY 904 CGCTCTCAACGGGGCGCTTTCTGTAGGGGCGAGAAATGTCATGACCCGACCGTCTCT 963  
DB 860 CTCCACTCAACGAGGGGCGCTTCTGCGAGGGCGAGCAATTCAGAA---GACCGCTGCA 916  
QY 964 CTCTGCTGCTCTGTGAGCGACCTGGAGCCCTGGAGCAAGTGGTGGCTGTGGGC 1023  
DB 917 CCACCATCTGCCAGTCAATGGGGCGGTGGACGAGTGGAGCAAGTGGTGGCTGTGAGCA 976  
QY 1024 TGGACTGCACCCACTGCGGGAGCGCTGTGAGTGTCTGTACCCAGCACCCCGACGAGGGG 1083  
DB 977 CTGAGTGTGCCACTGGCGTGTGCGCGAGTGTGATGGCGCCCGCCAGAACGAGGGCC 1036  
QY 1084 AGGAGTGCAGGGCACTGACCTGGACACCCCGCACTGTACCACTGACCTGTGTGTACACA 1143  
DB 1037 GTGACTGACGGGAGCGCTGTGCACTTAAGAACTGACAGATGGGTGTGTCATGCAAA 1096  
QY 1144 GTGCTTC-----TGCGCTGAGGAGCTGGGCC 1170  
DB 1097 GTGAGTACAGTGTGGTCTCTGTCCCGCAGTGTGAGGAGCTTCAGGGGATGCGCGC 1156  
QY 1171 TCTATGTGGGCTC---ATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTCA 1227  
DB 1157 TGTATGCGGGGCTGTGCTGTGCGCACTCTCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1216

QY	1228	TCTCGTTTATTGCCGGAAGAGGGGCTGGACTCAGATGTGGCTGACTCGTTCATT-	1286
Db	1217	TGGTGTGTATACCGCGCAACTGCCGTGACTTCGACACAGACATCACTGACTCATCTGCTG	1276
QY	1287	--CTCACTCAGGCTTCCAGCCCGTCAGCATCAAGCCGAGCAAGAGACACACCCCATC	1344
Db	1277	CCCTGACTGGTGTTCACCCCGTCAACTTTAAGACGCAAGGCCAGTAACCGCGACG	1336
QY	1345	TGCT-----CACCATCCAGCCGAGCTCAGCACACACCACTTACAGGCGAGTCTCT	1398
Db	1337	TCCTACACCCCTCTGTGCTCTGACTGACAGCCAGCCCGGCATCTACCGGGAACCG	1396
QY	1399	GTCCCGGAGGATGGCCCGAGCCCAAGTTCAGCTCACCAATGGGCACTGTCTCAGCC	1458
Db	1397	TGTATGCCCTGCAGGACTCACCGACAAATCCCATGACCAACTCTCTCTGCTGGACC	1456
QY	1459	CC-----	1460
Db	1457	CTTTACCCAGCCTTAAGGTCAAGTCTTACAGCTCCAGCACCAAGGCTCTGGGCCAGGCC	1511
QY	1461	-----CTGGGTGGGCGCGCCACACACTGTGACACCAAGCTCTCCCA	1500
Db	1517	TGGCAGATGGGCTGACCTCTGGGGTCTTGGCGCTGGCACATACCTTAGCGATTCTG	1576
QY	1501	CCTCTGAGGCGGAGGTTCTCTCCCGCTCTCCACCC-----AGAACTACTTCCGCT	1554
Db	1577	CCGGGACACCACTTCTCTGACCTGCGCAGCCAGCCTCGGTCCGACAGCTCTTGG	1636
QY	1555	CCCTGCCCCGAGGACACAGCAACATGACCTATGGGACCTTCACTTCTCTGGGGCCGGC	1614
Db	1637	GCCTGCCCCGAGACCCAGGGAGCAGCTCAGCGCACCTTTGGCTGCTCGGGTGGGAGCG	1696
QY	1615	TGATGATCCCTAATACAGGTATCAGCTCTCTCATCCCCCAGATGCCATACCCCGAGGGA	1674
Db	1697	TCAGATCCCCGGCACAGGTGTACGTTGTGTGTGCCAATGAGGCATTCGCCAGGCA	1756
QY	1675	AGATCTATGAGATACCTCAGCTGACAGCGCGGAAGACGTGAGTTGCCCTAGCTG	1734
Db	1757	AGTTCTACGAGATGATCTACTCATCAACAGGCAGAAAGTACCTCGCGCTTTCAGAAG	1816
QY	1735	GCTGTGACACCTGCTGAGTCCCATCGTTAGCTGTGACCCCTGGGGTCTGCTCACCC	1794
Db	1817	GGACCCAGACATATTGAGCCCTTCGGTGACCTGTGGACCCACAGGCTCTCTGTGCC	1876
QY	1795	GGCAGTCACTCCTGGCTATGACCACTGTGGGAGCCAGCCCTGCAGCTGAGGCTGC	1854
Db	1877	GCCCGCTCATCTCACCATGCCCACTGTGCCGAAGTCAGTGCCGTGACTGATCTTTC	1936
QY	1855	GCCTCAAAAGACGTCTGTGCGAGGGCAGCTGGGAGCAGGATGTCTGCACCTCGGGGAGG	1914
Db	1937	AGCTCAAGACCCAGGCCACAGGCCACTTGG--AGGAGGTGGTGACCTCGATGAGG	1993
QY	1915	AGGCGCCCTCCACCTCTACTCTGCAGCTGGAGGCGAGTGCCTGCTACTGCTTTCACCG	1974
Db	1994	AGACCTTGAACACACCTCTGCTACTGCGAGCTGGAGCCAGGGCTGTGCATCTCTGCTGG	2053
QY	1975	AGCAGCTGGCGCTTTTCCCTGTGTGGAGAGGCCCTCAGCGTGGCTGCCCGCCAAAGGCC	2034
Db	2054	ACCAGCTGGGACCTACGTGTTACGGCGAGTCTCTATTCCTCGCTCAGCAGTCAAGCGGC	2113
QY	2035	TCAAGCTGCTTGTGTTTGGCGGTGTGCTGCACTCTCTGTGAGTACAAATTCGGGTCT	2094
Db	2114	TCCAGCTGGCGCTCTTTCGCGCCCGCCCTCTGCACCTCTCTGGAGTACAGGCTCTCGGGTCT	2173
QY	2095	ACTGCTGATGACACCAACGATGCATCAAGGAGGTGTGCACCTGGAAGCAGCTGG	2154
Db	2174	ACTGCTGGAGACACGCTGTGACTGAAGGAGGTGCTGGAGCTGAGCGGACTCTGG	2233
QY	2155	GGGGACAGCTGATCAGGAGCCACGGGTCTCTGCATTCAGAGGACAGTTACCAACCTGC	2214
Db	2234	GCGGATACTTGTGTGGAGGAGCCGAACCGCTAATGTTCAAGGACAGTTACCAACCTGC	2293
QY	2215	GCCTATCCATCACGATGTGCCCAAGCTCCCTGTGTGAGAGTAAGCTCTCTGTCTAGCTACC	2274

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 12:08:07 ; Search time 197 Seconds  
(without alignments)  
8115.805 Million cell updates/sec

Title: US-09-970-944-1  
Perfect score: 2881  
Sequence: 1 agctgggctccgggtgag.....ccttccccacacggggaga 2881

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235.8	77.6	3014	2 US-08-808-982-1	Sequence 1, Appli
2	2235.8	77.6	3014	3 US-09-306-902A-1	Sequence 1, Appli
3	1570.4	54.5	1787	2 US-08-808-982-2	Sequence 2, Appli
4	1570.4	54.5	1787	3 US-09-306-902A-2	Sequence 2, Appli
5	1147.8	39.8	1282	4 US-09-833-381-1806	Sequence 1806, Ap
6	783.4	27.2	2831	2 US-08-808-982-3	Sequence 3, Appli
7	783.4	27.2	2831	3 US-09-306-902A-3	Sequence 3, Appli
8	419	14.5	1605	4 US-09-833-381-1807	Sequence 1807, Ap
9	256	8.9	771	1 US-08-253-155A-17	Sequence 17, Appl
10	119	4.1	305	2 US-08-808-982-4	Sequence 4, Appli
11	119	4.1	305	3 US-09-306-902A-4	Sequence 38, Appl
12	51.8	1.8	1280	3 US-08-483-533-38	Sequence 38, Appl
13	51.8	1.8	1280	4 US-09-283-471A-38	Sequence 38, Appl
14	51.2	1.8	699	4 US-09-252-991A-7947	Sequence 7947, Ap
15	51.2	1.8	1302	4 US-09-252-991A-7533	Sequence 7533, Ap
16	51.2	1.8	1947	4 US-09-252-991A-12265	Sequence 12265, A
17	50.8	1.8	1770	4 US-09-252-991A-12126	Sequence 12126, A
18	50.8	1.8	3885	4 US-09-369-364A-16	Sequence 16, Appl
19	50.8	1.8	4884	4 US-09-252-991A-12126	Sequence 12126, A
20	50.8	1.8	4884	4 US-09-252-991A-12292	Sequence 12292, A
21	50.2	1.7	1300	3 US-08-483-533-39	Sequence 39, Appl
22	50.2	1.7	1300	4 US-09-283-471A-39	Sequence 39, Appl
23	48.6	1.7	1068	4 US-09-252-991A-9933	Sequence 9933, Ap
24	48.6	1.7	1251	4 US-09-252-991A-10142	Sequence 10142, A
25	48	1.7	4257	2 US-08-690-473-1	Sequence 1, Appli
26	48	1.7	4257	3 US-09-259-821A-1	Sequence 1, Appli
27	48	1.7	4257	3 US-08-843-659-1	Sequence 1, Appli

c 28 48 1.7 8312 4 US-09-620-312D-1048 Sequence 1048, Ap  
c 29 47.8 1.7 1288 1 US-08-440-856A-9 Sequence 9, Appli  
c 30 47.8 1.7 2211 4 US-09-252-991A-11256 Sequence 11256, A  
31 47.8 1.7 2289 4 US-09-252-991A-10995 Sequence 10995, A  
32 47.6 1.7 1627 4 US-09-675-018B-11 Sequence 11, Appli  
33 47 1.6 7218 1 US-08-232-463-14 Sequence 14, Appli  
34 46.8 1.6 1951 3 US-08-922-865-1 Sequence 1, Appli  
35 46.8 1.6 1951 4 US-09-510-949-1 Sequence 1, Appli  
36 46.4 1.6 453 4 US-09-252-991A-3971 Sequence 3971, Ap  
37 46.4 1.6 567 4 US-09-252-991A-4004 Sequence 4004, Ap  
38 46.2 1.6 11220 3 US-09-105-537-32 Sequence 32, Appli  
39 46.2 1.6 36778 3 US-09-105-537-5 Sequence 5, Appli  
40 46.2 1.6 38506 3 US-09-320-878-19 Sequence 19, Appli  
41 46.2 1.6 38506 4 US-09-141-908-1 Sequence 1, Appli  
42 46.2 1.6 38506 4 US-09-657-440-19 Sequence 19, Appli  
43 46 1.6 1065 4 US-09-340-798A-31 Sequence 31, Appli  
44 46 1.6 3624 1 US-07-951-715A-6 Sequence 6, Appli  
45 46 1.6 3624 2 US-08-459-448A-6 Sequence 6, Appli

## ALIGNMENTS

RESULT 1  
US-08-808-982-1  
; Sequence 1, Application US/08808982  
; Patent No. 5939271  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsay  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,982  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3014 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; US-08-808-982-1

Query Match 77.6%; Score 2235.8; DB 2; Length 3014;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;  
QY 87 ATGGCGCTCCGGCGGCTGTGGCCAGCTCTCGGCATAGTCTCGCGCTTGGCTC 146  
|||||

Db 1 ATGGCGCTCCGGCCCGGCGCTGTGGCCAGTGTCTCTGGGCAATAGTCTCTCGCCGCTGGCTT 60  
QY 147 CCGCGCTCCGGTCCCGCAGCAGAGTGCCACCGTGGCCAAACCCAGTGCTCGTGGCCAAACCG 206  
Db 61 CGTGGTTCGGGTCCCGCAGCAGAGTGCCACCGTGGCCAAATCCAGTCCCGGTGCAACCC 120  
QY 207 GACCTGCTTCCCACTTCCTGGTGGAGCCGAGGATGTGTACATCTGTCAAGAACAGCCA 266  
Db 121 GACCTGCTGCCCACTTCCTGGTAGAGCTGTGAGGACGTGTACATTTGTCAAGAACAGCCG 180  
QY 267 GTGCTGCTTGTGGCAGGCGGTGGCCGCGCAGCAGATCTTCTCAAGTGCACAGCGGAG 326  
Db 181 GTGTTGTTGGTGGCAGGCTGTGCTGCCACCCAGATCTTCTCAAGTGCATGGGAA 240  
QY 327 TGGGTGGCCGAGGTGGACCACTGTATCGAGCGCAGCACAGCGGAGCAGTGTGTAGCCG 386  
Db 241 TGGGTCCGCCAGGTTCGATCAGCTAAATTGAACGCGACACCGACAGCAGCAGCGGATGCCA 300  
QY 387 ACCATGGAGTCCGCAATTAATGTCTCAGGCGCAGGTTCAGAGGTGTTCGGGCTGGAG 446  
Db 301 ACCATGGAGTCCGATATCAACGTATCGAGCGCAGGTAGAGAAAGTGTTCGGGCTGGAG 360  
QY 447 GATACGTGGTCCAGTCCGCTGGGATGAGTCTCTCGGCGACCAACCAAGAGTCAGAAAGCC 506  
Db 361 GAATACGTGGTCCAGTGTGGCATGGAGTCTCTCGGTACCCACCAAAAGTTCAGAGGCC 420  
QY 507 TACATCCGCATAGCCAGATTCGCAAGAACTTCGAGCAGAGCGCTGGCCAGAGGAGTG 566  
Db 421 TACATCCGATGCTTATTTGGCAAGAACTTTGAGCAGGAGCCACTGGGCCAAGGAAGTG 480  
QY 567 TCCTGGAGCAGGCGATCGTGTCTCCCTGCGTCCAGCGGCGATCCCTCCAGCCGAG 626  
Db 481 TCACTGAGCAAGCAATGTACTACCTGTGTGCCCCCAGAGGAATCCCCCAGCTGAG 540  
QY 627 GTGAGTGGTCCGAAAGAGGAGCTGTGTGACCGCTCCCTGGACCCCAATGTATATC 686  
Db 541 GTGAGTGGCTTCCAAATGAGACCTGTGTGACCCCTCCCTCGATCCCAATGTATATC 600  
QY 687 ACGGGAGCAGACACCTGTGTGTGCGACAGCCCGCTTGTGACAGCGCCAACTACACC 746  
Db 601 ACGGGAGCAGACACCTAGTGTGTGCGTCAAGCCCGCTGGCCGACAGCGCCAACTACACC 660  
QY 747 TCGTGGCCAAAGATCTGTGCGAGTGTGCGAGCGCTCCCTGCTGTCTCATCTCTAC 806  
Db 661 TGTGTGCCAAGAACTGTAGCCGTGCGGAGCACTCTGAGCGGTCAITGTTAT 720  
QY 807 GTGAACGTGGTGGTGTGCGACCGAGTGTCCGTCTGAGCGCCAGCTGTGGCGC 866  
Db 721 GTGAACGTGGTGTGCGACCGTGTGAGTGTGGTCCGTCTGAGCGCCAGCTGTGGCGT 780  
QY 867 GGCTGGCAGAAACGGAGCCGAGTGTGCAACCAACCGCGGCTCTCAACGGGGGCGCTTTC 926  
Db 781 GGCTGGCAGAAACGGAGCCGAGTGTGCAACCAACCGGCACTCTCAACGGGGGCGCTTTC 840  
QY 927 TGTGAGGGCAGATGTGCAATGACCGCAGCTCTCTCTCTGTGTGTGTGAGCGGC 986  
Db 841 TGTGAGGGCAGATGTGCAAGAAACAGCTGTGCGCACTCTG---TGGCCAGTGGATGG 897  
QY 987 AGCTGGAGCCCGTGGAGCAAGTGTGCGGCTGTGGGCTGTGAGCTGTGAGCTGTGGCGAGC 1046  
Db 898 AGCTGGAGTGTGAGTGTGAGTGTGAGCTGTGGGCTGTGAGCTGTGAGCTGTGGCGAGC 957  
QY 1047 CGTGAAGTGTGACCCAGCAGCCCGCAACGGAGGGGAGGTGCGAGGCACTGACCTG 1106  
Db 958 CGGAGTGTCTGACCCAGCAGCCCGCAATGAGGTGAGGAGTGTGCGGGGTGCTACCTG 1017  
QY 1107 GACACCGCACTGTACAGTACCTCTGTGTACACAGTGTCTGTGGCCCTGAGGACGTG 1166  
Db 1018 GACACCGCACTGTACAGTACCTCTGTGTGTGCAACCGCTTCTGTGCCCGGAGAGCTG 1077  
QY 1167 GCGCTCTATGTGGGCGCTATCGCCGCTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1226  
Db 1078 GCTCTCTACATCGGCTTGTGCTGTGGCTGTGTGCTTCTTGTGTGTGCTGTGCTT 1137

QY 1227 ATCTCTGTTTATTGCGGAAGAGGAGGGGCTGGACTCAGATGTGGCTGACTGCTGCTATT 1286  
Db 1138 GGACTCATTTATCTGTGCAAGAGGAGGGCTGGACTCCGATGTGGCGACTCGTCCATC 1197  
QY 1287 CTCACCTCAGGCTTCCAGCCCTCAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAG 1346  
Db 1198 CTCACCTCGGCTTCCAGCTGTGAGCATCAAGCCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1257  
QY 1347 CTCACCATCAGCGGAGCTCAGCACCAGC---ACCACCTACCGAGGAGTCTCTGCTCC 1403  
Db 1258 CTCACCATCAGCGGAGCTCAGCACCAGCCTACCACTACCACTACCACTACCACTACCACTAC 1317  
QY 1404 CGGAGATGGGCGGCGGCGGCTTCCAGCTCACCAATGGGCACCTGTTCAGCCCCCTG 1463  
Db 1318 AGGAGATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1377  
QY 1464 GGTGGCGGCGGCGACACATGACACACAGTCTTCCACCTCTGAGCGGAGGATGCTGTC 1523  
Db 1378 GGGAGTGGCGGCGCATACGTTGACACACAGCTCACCCACCTCTGAGGTGAGGACTTCGTC 1437  
QY 1524 TCCGGCTCTCCACCCAGAACCTACTTCCGCTCCCTGCGCGGCGGCGGCGGCGGCGGCGGCGG 1583  
Db 1438 TCCGGCTCTCCACCCAGAACCTACTTTCGTTTCCCTGCGCGGCGGCGGCGGCGGCGGCGGCGG 1497  
QY 1584 TATGGGACCTTCAACTTCTCGGGGCGGCGCTGATGATCCCTTAATACAGGTATACGCTC 1643  
Db 1498 TACGGGACCTTCAACTTCTCGGGGCGGCGCTGATGATCCCTTAATACGGGGATCAGGCTC 1557  
QY 1644 CTATCCCGGCGGAGTGCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1703  
Db 1558 CTATACCCCGGAGTGCATACCCCGGAGGGAAGATCTACGAGATCTACCTCAGCTGCAC 1617  
QY 1704 AAGCGGAGACCTGAGTGTGCGCTAGCTGCTGTCAGACCCCTGCTGAGTTCGCTCATCGTT 1763  
Db 1618 AAGCGGAGACCTGAGTGTGCGCTAGCTGCTGTCAGACCCCTGCTGAGTTCGCTCATCGTT 1677  
QY 1764 AGCTGTGGAACCCCTGCGCTCTGCTCACCGGCGGAGTCACTCTGCTGATGGACCACTGT 1823  
Db 1678 AGCTGTGGGCGGCGGAGTGTGCTGCTCACCGGCGGAGTCACTCTGCAATGGACCACTGT 1737  
QY 1824 GGGAGCCGCGGCTGACAGTGGAGCTGCGCTCAAAAGAGAGTGTGCGAGGGGAGC 1883  
Db 1738 GGAGAGCCGCGCTGACAGTGGAGTCTGCGCTCAAAAGAGAGTCTGCGAGGGGAGT 1797  
QY 1884 TGGAGCAGATGCTGACCTGCGGAGGCGGCTCCACCTCTACTACTGCGCAG 1943  
Db 1798 TGGG---AGGATGCTGCACTTGGTGGAGTCACTTCCACCTCTACTACTGCGCAG 1854  
QY 1944 CTGAGGCGCAGTGTGCTGCTACTCTTCAAGAGAGTGGGCGGCTTTCCTGCTGGTGGGA 2003  
Db 1855 CTGAGGCGGCGGCTGCTATGCTTCAAGAGAGTGGGCGGCTTTCCTGCTGGTGGGA 1914  
QY 2004 GAGGCGCTCAGCTGCTGCGGCGGAGGCGGCTCAAGCTGCTTCTGCTGCTGCTGCTGCTGCT 2063  
Db 1915 GAGGCGCTCAGCTGCTGCGGCGGAGGCGGCTCAAGCTGCTTCTGCTGCTGCTGCTGCTGCT 1974  
QY 2064 TGCACCTCCCTCAGTACCAATCCGGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2123  
Db 1975 TGTAGTCCCTTGTAGTACCAATCCGAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2034  
QY 2124 AAGGAGTGTGTCAGCTGAGGAAGAGCTGGGGGAGAGTGTGATCCAGGAGCAGGGTGC 2183  
Db 2035 AAGGAGTGTGTCAGCTGAGGAAGAGCTAGTGGACAGTGTGATCCAGGAGCCTCGCTGC 2094  
QY 2184 CTGCACTTCAAGGACAGTACCAACCTGCGCTTATCCATCCAGGATGCTGCGCAGCTCC 2243  
Db 2095 CTGCACTTCAAGGACAGTACCAACCTAGCTCTCTCATCCAGAGCTGCTGCTGCTGCTGCTGCTGCT 2154  
QY 2244 CTGTGGAAGAGTAACTCTTGTAGCTACAGGAGATCCCTTTTATCATCTGGAAT 2303  
Db 2155 CTGTGGAAGAGCAAGCTACTTGTAGCTACAGGAGATCCCTTTTACCACATCTGGAAC 2214





QY	927	TTGTGAGGGGCGAGATGTTCATGACCGCAGCTTCTCTCTCTGTTGTCTCTGTGGACGGC	986
Db	841	TTGTGAGGGGCGAGATGTTCAGAAAAAGAGCTTGGCCACTCTG---TGCCAGTGGATGGG	897
QY	987	AGCTGAGCCCGTGGAGAGTGTCTGGCTGTGGCTGTGACTGTGACCACTGCGGGAGC	1046
Db	898	AGCTGAGTTCGTGGAGTGAAGTGTGAGCTGTGGCTGTGACTGTGACCACTGCGGGAGC	957
QY	1047	CGTGAAGTGTCTGACCCAGCAGCCCGCAAGCGAGGGGAGAGTGCAGGGCACTGACCTG	1106
Db	958	CGCGAGTGTCTGACCCAGCAGCCCGCAATGGAGGTGAGAGTGTGCGGGTGTGACCTG	1017
QY	1107	GACACCGGAACTGTACAGTGAACCTCTGTGTACAGTGTCTTGTGGCTGTGAGGACGTG	1166
Db	1018	GACACCGGAACTGTACAGTGAACCTCTGTGTACAGTGTCTTGTGGCTGTGAGGACGTG	1077
QY	1167	GCCTCTATGTGGGCTCATGCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGCTC	1226
Db	1078	GCTCTCTATGCTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGCTC	1137
QY	1227	ATCCTCGTTTATTTCCCGAAGAGAGGGGCTGGAATCAGATGTGGCTGACTGTGCTGAT	1286
Db	1138	GGACTCATTTACTGTGCGAAGAGAGAGGGCTGGAATCAGATGTGGCTGACTGTGCTGAT	1197
QY	1287	CTCAGCTCAGGCTTCAGCCCGCTCAGATCAAGCCAGCAAGAGAGAGAGAGAGAGAGAG	1346
Db	1198	CTCAGCTCAGGCTTCAGCCCGCTCAGATCAAGCCAGCAAGAGAGAGAGAGAGAGAGAG	1257
QY	1347	CTCAGCTCAGGCTTCAGCCCGCTCAGATCAAGCCAGCAAGAGAGAGAGAGAGAGAGAG	1403
Db	1258	CTCAGCTCAGGCTTCAGCCCGCTCAGATCAAGCCAGCAAGAGAGAGAGAGAGAGAGAG	1317
QY	1404	CGGAGGATGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1463
Db	1318	AGGAGGATGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1377
QY	1464	GGTGGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1523
Db	1378	GGGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1437
QY	1524	TCCGCGCTCTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1583
Db	1438	TCCGCGCTCTCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1497
QY	1584	TATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTTAATGATGATGATGATGATGAT	1643
Db	1498	TACGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTTAATGATGATGATGATGATGAT	1557
QY	1644	CTCATCCCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1703
Db	1558	CTCATCCCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1617
QY	1704	AAGCGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1763
Db	1618	AAGCGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1677
QY	1764	AGCTGTGACCCCTGTGGGCTCTGTCTACCGGCGAGTATCTGCTGCTGCTGCTGCTGCTGCT	1823
Db	1678	AGCTGTGACCCCTGTGGGCTCTGTCTACCGGCGAGTATCTGCTGCTGCTGCTGCTGCTGCT	1737
QY	1824	GGGAGCCCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	1883
Db	1738	GGGAGCCCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT	1797
QY	1884	TGGAGCAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1943
Db	1798	TGGG---AGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1854
QY	1944	CTGAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2003
Db	1855	CTGAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1914
QY	2004	GAGGCGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2063

RESULT 3  
US-08-808-982-2  
; Sequence 2, Application US/08808982  
; Patent No. 5919271  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsey  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:



ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/808,982  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UC96-217  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1787 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-808-982-2

Query Match 54.5%; Score 1570.4; DB 2; Length 1787;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1721; Conservative 0; Mismatches 36; Indels 19; Gaps 11;

QY 1114 GCAACTGTACCACTGACCTCTGTGTACAGTCTTCTGSCCTGAGGAGCTGGCCCTCT 1173  
DB 1 GCAACTGTACCACTGACCTCTGTGTACAGTCTTCTGSCCTGAGGAGCTGGCCCTCT 59  
QY 1174 ATGTGGGCTCATGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGT 1233  
DB 60 ATGTGGGCTCATGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGT 119  
QY 1234 TTTATTTGCGGAGAGAGAGGGGCTGAGTCTGAGTGTGGCTGACTGCTCATTTCTCACT 1293  
DB 120 TTTATTTGCGGAGAGAGAGGGGCTGAGTCTGAGTGTGGCTGACTGCTCATTTCTCACT 179  
QY 1294 CAGGCTTCCAGCCGCTCAGCATC-AAGCCAGAAAGAGAGAAACCCCATCTGTCTACC 1352  
DB 180 CAGGCTTCCAGCCGCTCAGCATCTAAGCCAGAAAGAGAGAAACCCCATCTGTCTACC 239  
QY 1353 ATCCAGCCGAGCTCAG---CACCACACCACTACAGGGGAGTCTGTCCCGGAG 1409  
DB 240 ATCCAGCCGAGCTCAGCAGACACCACTACAGGGGAGTCTGTCCCGGAG 299  
QY 1410 GATGGGCGGAGCCCAAGTTTCAGCTCAGAAATGGGCACTGTCTCAGCCCTGTGGTGC 1469  
DB 300 GATGGGCGGAGCCCAAGTTTCAGCTCAGAAATGGGCACTGTCTCAGCCCTGTGGTGC 359  
QY 1470 GSCCGCCACACACTGACCAAGCTCTCCCACTCTGAGGCGGAGAGTGTCTCCGCG 1529  
DB 360 GSCCGCCACACACTGACCAAGCTCTCCCACTCTGAGGCGGAGAGTGTCTCCGCG 419  
QY 1530 CTCTCCACCCAGAACTACTTCCGCTCCCTGCGCCGAGGACCAAGATGATGATGATG 1589  
DB 420 CTCTCCACCCAGAACTACTTCCGCTCCCTGCGCCGAGGACCAAGATGATGATGATG 479  
QY 1590 ACCTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATC 1649  
DB 480 ACCTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATC 539  
QY 1650 CCCCAGATGCCATACCCGAGGAGGAGATCTATGATCTACTCTCAGCTGTGCAAGCGG 1709  
DB 1612 GCTGGGAGTGGCCAGGAGCGGTGGCTTTTTCACAGTGTTCGGAGGCTGAGTGTGAG 1671

DB 540 CCCCAGATGCCATACCCCGGAGGAGATCTATGAGATCTACCTCAGCTGCAAGCGG 599  
QY 1710 GAAGACGTGAGTTGCCCTAGCTAGCTGTGTGAGACCTCTGTGAGTCCCATCTGTAGCTGT 1769  
DB 600 GAAGACGTGAGTTGCCCTAGCTAGCTGTGTGAGACCTCTGTGAGTCCCATCTGTAGCTGT 659  
QY 1770 GGACCCCTCGGCTGCTGCTCAGCCGCGAGTCTCTCTGCTATGAGACCTGTGGGAG 1829  
DB 660 GGACCCCTCGGCTGCTGCTCAGCCGCGAGTCTCTCTGCTATGAGACCTGTGGGAG 719  
QY 1830 CCGAGCCCTGACAGCTGAGCTGCGCTCAAAAAGCAGTCTGCGAGGAGGAGCTGGAG 1889  
DB 720 CCGAGCCCTGACAGCTGAGCTGCGCTCAAAAAGCAGTCTGCGAGGAGGAGCTGGAG 778  
QY 1890 CAGGATGTGTGCACTTGGGCGAGGAGGCGCTCCACCTCTACTACTGCGAGCTGGAG 1949  
DB 779 GATG---TCTGCACTGTGGCGAGGAGGCGCTCCACCTCTACTACTGCGAGCTGGAG 834  
QY 1950 GCCAGTGCCTGCTACGTCTTTCAGGAGCAGCTGGGCGCTTTCCTGTGGGAGGAGGCC 2009  
DB 835 GCCAGTGCCTGCTACGTCTTTCAGGAGCAGCTGGGCGCTTTCCTGTGGGAGGAGGCC 894  
QY 2010 CTCAGCTGCTGCTGCGCCAGGCGCTCAAGCTGCTTCTGTTTGGCGGCTGCGTGCACC 2069  
DB 895 CTCAGCTGCTGCTGCGCCAGGCGCTCAAGCTGCTTCTGTTTGGCGGCTGCGTGCACC 954  
QY 2070 TCCCTCAGTACAAATCCGGGTCTACTGCTGCATGACACCCACGATGCATCAAGGAG 2129  
DB 955 TCCCTCAGTACAAATCCGGGTCTACTGCTGCATGACACCCACGATGCATCAAGGAG 1014  
QY 2130 GTGGTGCAGCTGGAAGAGCAGCTGGGCGGAGCAGCTGATCCAGGAGCAGGGTCTCTGCAC 2189  
DB 1015 GTGGTGCAGCTGGAAGAGCAGCTGGGCGGAGCAGCTGATCCAGGAGCAGGGTCTCTGCAC 1074  
QY 2190 TTCAAGGAGAGTTTACCAACCTGCGCTTATCCATCCAGAGTGTGCCAGCTCCCTGTGG 2249  
DB 1075 TT-AAGGAGAGTTTACCAACCT--GCCCTATCATCCAGATGTGCCAGCTCCCTGTGG 1131  
QY 2250 AAGATGAAGCTCTGTGTGAGTACCAAGGAGTCCCTTTTATCAGATCTGGAATGACG 2309  
DB 1132 AAGATGAAGCTCTGTGTGAGTACCAAGGAGTCCCTTTTATCAGATCTGGAATGACG 1191  
QY 2310 CAGCGGTACTTGCATGCACTTCCCTTGAGGAGTGTGAGCCCGAGCTAGTACCTG 2369  
DB 1192 CAGCGGTACTTGCATGCACTTCCCTTGAGGAGTGTGAGCCCGAGCTAGTACCTG 1251  
QY 2370 GCTTGAAGCTGTGGGTGTGGAGGCGAGCGGAGGAGCTTCAAGCATCAACTTC 2429  
DB 1252 GCTTGAAGCTGTGGGTGTGGAGGCGAGCGGAGGAGCTTCAAGCATCAACTTC 1311  
QY 2430 AACATCAAGGAGCAAGGTTTGTGAGTGTGGCTCTGGAGAGTGAAGCGGGGTC 2489  
DB 1312 AACATCAAGGAGCAAGGTTTGTGAGTGTGGCTCTGGAGAGTGAAGCGGGGTC 1371  
QY 2490 CCAGCCCTGTGGGCGGCGGAGTGTGAGTGTGGCTCTCAATTCGGGAGAGATAAT 2549  
DB 1372 CCAGCCCTGTGGGCGGCGGAGTGTGAGTGTGGCTCTCAATTCGGGAGAGATAAT 1431  
QY 2550 TCCAGCTTGGAGCCCACTGTAGGCGGCGTGCAGTGTGGGAGTGTGGGAGGAGTGAAG 2609  
DB 1432 TCCAGCTTGGAGCCCACTGTAGGCGGCGTGCAGTGTGGGAGTGTGGGAGGAGTGAAG 1491  
QY 2610 CACTTGAAGCAAGGAGTGTGAGTGTGGCTCTCAAGTGTGGGAGTGTGGGAGGAGTGAAG 2669  
DB 1492 CACTTGAAGCAAGGAGTGTGAGTGTGGCTCTCAAGTGTGGGAGTGTGGGAGGAGTGAAG 1551  
QY 2670 AACCTGTGGGAGGCGGCGGAGTGTGAGTGTGGCTCTCAAGTGTGGGAGTGTGGGAGG 2729  
DB 1552 AACCTGTGGGAGGCGGCGGAGTGTGAGTGTGGCTCTCAAGTGTGGGAGTGTGGGAGG 1611  
QY 2730 GCTGAGCTGGGCGGAGGAGGAGTGTGGCTCTCAAGTGTGGGAGTGTGGGAGGAGTGAAG 2787  
DB 1612 GCTGAGCTGGGCGGAGGAGGAGTGTGGCTCTCAAGTGTGGGAGTGTGGGAGGAGTGAAG 1671

QY	2788	GGCGGCGGAGCCCGACCTACCTACCTCTACAGCTTTGGCA--CCACCAAGGACAGGCA	2845	Db	180	CAGGCTTCAGCCCGCTAGCATCTAAGCCCGACGAAAGACAGACAACCCCATCTGCTCAAC	239
Db	1672	GGCGGCGGAG--CGAACACTACAAATTTTACAGTTTGGGAACCCACCAAGGACAGGCA	1729	QY	1353	ATCCAGCCGAGCTCAG---GACCAACCACCACTACAGGCGAGTCTCTGTCCCGGCGAG	1409
QY	2846	GAAGCGGAGCAGGCGCTTCCCGACACCGGGGAGA	2881	Db	240	ATCCAGCCGAGCTCAGCAACCACTACAGGCGAGTCTCTGTCCCGGCGAG	239
Db	1730	GAAGCGGAGCAGGCGCTTTCACCAAAACCGGGGAGA	1765	QY	1410	GATGGGCCAGCCCGCAAGTTCCAGCTCACCATATGGGCACCTGCTCAGCCCCCTGGGTGGC	1469
RESULT 4							
US-09-306-902A-2							
; Sequence 2, Application US/09306902A							
; Patent No. 6277585							
; GENERAL INFORMATION:							
; APPLICANT: Tessier-Lavigne, Marc							
; Leonardo, E. David							
; Hink, Lindsey							
; Masu, Masayuki							
; Kazuko, Keino-Masu							
; TITLE OF INVENTION: Netrin Receptors							
; NUMBER OF SEQUENCES: 9							
; CORRESPONDENCE ADDRESS:							
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP							
; STREET: 268 BUSH STREET, SUITE 3200							
; CITY: SAN FRANCISCO							
; STATE: CALIFORNIA							
; COUNTRY: USA							
; ZIP: 94104							
; COMPUTER READABLE FORM:							
; MEDIUM TYPE: Floppy disk							
; COMPUTER: IBM PC compatible							
; OPERATING SYSTEM: PC-DOS/MS-DOS							
; SOFTWARE: PatentIn Release #1.0, Version #1.30							
; CURRENT APPLICATION DATA:							
; APPLICATION NUMBER: US/09/306,902A							
; FILING DATE: 07-May-1999							
; CLASSIFICATION: <Unknown>							
; ATTORNEY/AGENT INFORMATION:							
; NAME: OSMAN, RICHARD A							
; REGISTRATION NUMBER: 36,627							
; REFERENCE/DOCKET NUMBER: UC96-217							
; TELECOMMUNICATION INFORMATION:							
; TELEPHONE: (415) 343-4341							
; TELEFAX: (415) 343-4342							
; INFORMATION FOR SEQ ID NO: 2:							
; SEQUENCE CHARACTERISTICS:							
; LENGTH: 1787 base pairs							
; TYPE: nucleic acid							
; STRANDEDNESS: double							
; TOPOLOGY: linear							
; MOLECULE TYPE: cDNA							
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:							
US-09-306-902A-2							
Query Match 54.5%; Score 1570.4; DB 3; Length 1787;							
Best Local Similarity 96.9%; Pred. No. 0;							
Matches 1721; Conservative 0; Mismatches 36; Indels 19; Gaps 11;							
QY	1114	GCAACTGTACAGTAGACCTCTGTGTACACAGTCTCTGGCCCTGAGGACGTGGCCCTCT	1173	QY	1114	GCAACTGTACAGTAGACCTCTGTGTACACAGTCTCTGGCCCTGAGGACGTGGCCCTCT	1173
Db	1	GCAACTGTACAGTAGACCTCTGTGTACACAGTCTCTGGCCCTGAGGACGTGGCCCTCT	59	Db	1	GCAACTGTACAGTAGACCTCTGTGTACACAGTCTCTGGCCCTGAGGACGTGGCCCTCT	59
QY	1174	ATGTGGGCTCATCGCCGTGGCCGTTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1233	QY	1174	ATGTGGGCTCATCGCCGTGGCCGTTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1233
Db	60	ATGTGGGCTCATCGCCGTGGCCGTTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	119	Db	60	ATGTGGGCTCATCGCCGTGGCCGTTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	119
QY	1234	TTTATTGCGGAAGAGGAGGCGCTGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1293	QY	1234	TTTATTGCGGAAGAGGAGGCGCTGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1293
Db	120	TTTATTGCGGAAGAGGAGGCGCTGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	179	Db	120	TTTATTGCGGAAGAGGAGGCGCTGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	179
QY	1294	CAGGCTTCAGCCCGTACATC-AAGCCAGCAAGCAGACAACCCCATCTGCTCACC	1352	QY	1294	CAGGCTTCAGCCCGTACATC-AAGCCAGCAAGCAGACAACCCCATCTGCTCACC	1352

QY	2430	AACATCACCAAGGACACAAGGTTTGTGAGTGTCTGGAGTGAAGCGGGGTC	2489
Db	1312	AACATCACCAAGGACACAAGGTTTGTGAGTGTCTGGAGTGAAGCGGGGTC	1371
QY	2490	CCAGCCCTGTGTGGGCCCCAGTGCCTTCAAGATCCGCTTCCTCATTTCCGACGAAGATATT	2549
Db	1372	CCAGCCCTGTGTGGGCCCCAGTGCCTTCAAGATCCGCTTCCTCATTTCCGACGAAGATATT	1431
QY	2550	TCCAGCCTTGGACCCACACCTGTAGCGGGGGTCCGACTTGGCGGACTCTGGCCCCAGAAATC	2609
Db	1432	TCCAGCCTTGGACCCACACCTGTAGCGGGGGTCCGACTTGGCGGACTCTGGCCCCAGAAATC	1491
QY	2610	CACCTGGACAGCCCATCTCAGCTTTCTTTCCTCCAAAGCCAGCCCCACAGCCATGATCCTC	2669
Db	1492	CACCTGGACAGCCCATCTCAGCTTTCTTTCCTCCAAAGCCAGCCCCACAGCCATGATCCTC	1551
QY	2670	AACCTGTGGGAGGCGCGGCATTTCCCCAAACGGCAACCTCAGCCAGCTTGGCTGCAGCAGTG	2729
Db	1552	AACCTGTGGGAGGCGCGGCATTTCCCCAAACGGCAACCTCAGCCAGCTTGGCTGCAGCAGTG	1611
QY	2730	GCTGGACTGGGCGCAGCAGACGTCGGGCTC-TTCACAGTG-TGGAGGCTGAGTGTCTGAG	2787
Db	1612	GCTGGACTGGGCGCAGCAGACGTCGGGCTC-TTCACAGTG-TGGAGGCTGAGTGTCTGAG	1671
QY	2788	GCGGGCCAGCGCCCGACACCTTACACTCTCACAGCTTTGGCA--CCCAACCAAGCAGGCA	2845
Db	1672	GCGGGCCAGG--CGAACACTACAAATTACAGTTTTTGGGAACCCCAAGGACAGGCA	1729
QY	2846	GAAGCCGGACAGGGGGCCCTTTCCCCACACCGGGGAGA	2881
Db	1730	GAAGCCGGACAGGGGGTTTTTCCCAAAACCGGGGAGA	1765

## RESULT 5

US-09-833-381-1806  
; Sequence 1806, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: NO. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE C...

```

; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

```

Query Match	Score 1147.8; DB 4;	Length 1282;
Best Local Similarity	39.8%;	
	97.7%;	Pred. No. 3.3e-240;

	Matches	1259;	Conservative	0;	Mismatches	18;	Indels	11;	Gaps	9;
QY	1510	CGAGAGAGTTCGTCTCCGGCTCTCCACCAGAACTACTTCGGTCCCTGCGCCGAGGCA	1569							
Db	1	CGAGAGAGTTCGTCTCCGGCTCTCCACCAGAACTACTTCGGTCCCTGCGCCGAGGCA	60							
QY	1570	CGAGCAACATGACCTATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTTAATA	1629							
Db	61	CGAGCAACATGACCTATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTTAATA	120							
QY	1630	CAGGTATCAGCTCTTCATCCCCCGAGTCCCATACCCGAGGGAAGATCTATGAGATCT	1689							
Db	121	CAGGAATCAGCTTCCTCATCCCCCGAGTCCCATACCCGAGGGAAGATCTATGAGATCT	180							

QY	1690	ACCTCAGCTGCACAAGCCGGAAGACGTGAGTGGTCCCTAGCTGGGTGTGAGACCTTCG	1749
Db	181	ACCTCAGCTGCACAAGCCGGAAGACGTGAGTGGTGGCCCTAGCTGGGTGTGAGACCTTCG	240
QY	1750	TGAGTCCCATGTTAGCTGTGGACCCCT--GGGTCTCTGTCTCACCCGGCCAGTCACTCT--	1807
Db	241	TGAGTCCCATGTTAGCTGTGGACCCCTGGGGCTCTGTCTCACCCGGCCAGTCACTCTG	300
QY	1808	GGCTATGACCACTGTGTGGGAGCCAGCCCTGCAGCT--GGAGCTTGGGCTCAAAAAGC	1866
Db	301	GGGTATGACCACTGTGTGGGAGCCAGCCCTGCAGCTGGAGGCTCGGCTCAAAAAGC	360
QY	1867	AGTCGTGAGGGCAGCTGGAGCAGGATGTGTCTGCACCTGGGCGAGAGCGCGCTCC	1926
Db	361	AGTCGTGAGGGCAGCTGGAGGATGTGTCT--TGCACCTGGGCGAGAGCGCGCTCC	418
QY	1927	ACCTCTACTCTGCCAGCTGGAGGCACTGGCTGCTCTAGCTTCTACCGAGCAGCTGGGCC	1986
Db	419	ACCTCTACTCTGCCAGCTGGAGGCACTGGCTGCTCTAGCTTCTACCGAGCAGCTGAGCC	478
QY	1987	GCTTTGCCCTGGTGGGAGAGGCCCTCAGCTGGCTGCCGCAAGCGGCTCAAGCTGCTTC	2046
Db	479	GCTATGCCCTGGTGGGAGAGGCCCTCAGCTGGCTGCCGCAAGCGGCTCAAGCTGCTTC	538
QY	2047	TGTTTGGCCGGTGGCTCTGCACCTCCTCGAGTACAACATCCGGGTCTACTGCTCTGCATG	2106
Db	539	TGTTTGGCCGGTGGCTCTGCACCTCCTCGAGTACAACATCTTGCTTCTACTGCTCTGCATG	598
QY	2107	ACACCCAGATGCACTCAAGGAGTGGTCAGCTGGAGAAGCAAGCTGGGGGACAGCTGA	2166
Db	599	ACATCAGATGCACTCAAGCTAGTGGTCAGCTGGAGAAGCAAGCTGGGGGACAGCTGA	658
QY	2167	TCCAGGAGCCAAGGCTCCTCACTTCAAGGACAGTTTACCACACCTGCGGCTATCCATCC	2226
Db	659	TCCAGGAGCCAAGGCTCCTCACTTCAAGGACAGTTTACCACACCTGCGGCTATCCATCC	718
QY	2227	ACGATGCCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTGAGCTACCAAGGAGATCCCT	2286
Db	719	ACGATGCCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTGAGCTACCAAGGAGATCCCT	778
QY	2287	TTTATCACATCTGGATGGCACCCAGGGTACTTGCATCTGCACTTCACTTCACTTCACTT	2346
Db	779	TTTATCACATCTGGATGGCACCCAGGGTACTTGCATCTGCACTTCACTTCACTTCACTT	838
QY	2347	TCAGCCCCAGCACTAGTGACCTGGCTGCAAGCTGGGTGGGAGTGGAGGGGAGC	2406
Db	839	TCAGCCCCAGCACTAGTGACCTGGCTGCAAGCTGGGTGGGAGTGGAGGGGAGC	898
QY	2407	GGCAGAGCTTACAGTCAACTTCAACATCACAAGACACAAGGTTTGTGAGCTGTCTGG	2466
Db	899	GGCAGAGCTTACAGTCAACTTCAACATCACAAGACACAAGGTTTGTGAGCTGTCTGG	958
QY	2467	CTCTGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCCCAAGTGCCTTCAAGATCCCT	2526
Db	959	CTCTGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCCCAAGTGCCTTCAAGATCCCT	1018
QY	2527	TCCTCATTCGGCAGAGATAA--TTTCCAGCCTGGACCCACCTGTAGCGGGGTGCCAC	2585
Db	1019	TCCTCATTCGGCAGAGATAA--TTTCCAGCCTGGACCCACCTGTAGCGGGGTGCCAC	1078
QY	2586	TGSCGG--ACTCTGGCCCCGAAACTCCACTGGACAGCCATCTCAGCTTCTTTGCTCCAA	2644
Db	1079	TGSCGGAACCTTGGCCCCGAAACTCCACTGGACAGCCATCTCAGCTTCTTTGCTCCAA	1138
QY	2645	GCCAGCCCCCAGCGATGATCTCTCAACTGTGGAGCGCGGCACTTCCCAACGGCAA	2704
Db	1139	GCCAGCCCCCAGCGATGATCTCTCAACTGTGGAGCGCGGCACTTCCCAACGGCAA	1198
QY	2705	CCTCAGCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGCAGCTGGCTCTTAC	2764
Db	1199	CCTCAGCAGCTGG--TGACAGAGTGGCTGGACTGGGC--AGCAGACGCTGGCTCTTAC	1254

QY	2765	ACTGTCGAGGCTGAGTGTGAGGCGG	2792
Db	1255	ACTGTCGAGGCTGAGTGTGAGGCGG	1282
RESULT 6			
US-08-982-3			
; Sequence 3, Application US/0808982			
; Patent No. 5939271			
; GENERAL INFORMATION:			
; APPLICANT: Tessier-Lavigne, Marc			
; APPLICANT: Leonardo, E. David			
; APPLICANT: Hink, Lindsay			
; APPLICANT: Masu, Masayuki			
; APPLICANT: Kazuko, Keino-Masu			
; TITLE OF INVENTION: Netrin Receptors			
; NUMBER OF SEQUENCES: 8			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP			
; STREET: 268 BUSH STREET, SUITE 3200			
; CITY: SAN FRANCISCO			
; STATE: CALIFORNIA			
; COUNTRY: USA			
; ZIP: 94104			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA: US/08/808,982			
; FILING DATE:			
; CLASSIFICATION: 530			
; ATTORNEY/AGENT INFORMATION:			
; NAME: OSMAN, RICHARD A			
; REGISTRATION NUMBER: 36,627			
; REFERENCE/DOCKET NUMBER: UC96-217			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (415) 343-4341			
; TELEFAX: (415) 343-4342			
; INFORMATION FOR SEQ ID NO: 3:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2831 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: cdna			
US-08-982-3			
Query Match 27.2%; Score 783.4; DB 2; Length 2831;			
Best Local Similarity 59.4%; Pred. No. 5.7e-161;			
Matches 1625; Conservative 0; Mismatches 971; Indels 139; Gaps 12;			
QY	184	ACCAGTGCCTGGGCGCAACCGGACCTGCTTCCCACTTCTGCTGGAGCCCGAGGATG	243
Db	104	ACTCCTTCCCATAGCACCCCGGAGCTGCTCCTTCTGCTGAAACGAGGATG	163
QY	244	TGTATCATCGTCAAGAACCAAGCCAGTGTCTTGTGTGCAAGGCCGTGCCCGCAGCAGA	303
Db	164	CCTACATCGTAAGAACCAAGCCAGTGTGAAATGCACTGCCGAGCCTTCCCTGCCACACAGA	223
QY	304	TCTTCTCAAGTGCAACCGGGAGTGGGTGGCCAGTGACCACTGATCGAGCGCAGA	363
Db	224	TCTACTTCAAGTGAATGCGAGTGGGTAGCCAGAAAGGCCACGTACCGCAGGAGGCC	283
QY	364	CAGACGGGACAGTGTGTGAGCGGACCATGAGGTCGGCATTAATGTCTCAAGGCGACAGG	423
Db	284	TGATGAGGCCACAGGCTTGGCATACAGAGGTGCCATAGAGTGTGCGGCGACGAGG	343
QY	424	TCGAGAAGTGTTCGGGTGGAGGAATATCTGGTGCAGTGGCGTGGCATGGAGCTCTCGG	483
Db	344	TGGAGGAATTTTGGGGTCGAGGACTACTGTGTGTCAGTGGCTGGGCTGTCTCGG	403
QY	543	GCACCACCAAGAGTCAGAAAGGCTCATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC	543
Db	463	GAAACCAACAGAGTCGCCGAGCTTACATCCGATTGCTTACCTTGGCAAGAACTTTGACC	463
QY	544	AGGAGCCGCTGGCCAAAGAGGTTCCCTGGAGAGGGCATCTGTCTGCCCTTGCCTGCCAC	603
Db	464	AGGAGCCCTGTGGCAAGAGGTTACCTTGGATCATGAGTCCCTTCTGCAAGTGGCCGAC	523
QY	604	CGGAGGGCATCCTCCAGCCGAGGTGGAGTCCGGAACGAGGACCTTGGTGGACCCGT	663
Db	524	CAGAGGGAGTGCCTGTGGCTGAGGTGGAATGGCTCAAGATGAAGATGATCATCATCCCG	583
QY	664	CCCTGACCCCAATGATATATCATCACGCGGAGACACAGCCCTGTGTGGCAGAGCCGCGC	723
Db	584	CTCAGAGCACTAATCTCTCTGCTTCCCTGCTTACCATTTGACCAACCTCATCTCCGCGC	643
QY	724	TTGCTGACACGGCCAACTACCTGCTGGCGGCAAGAACTATCTGTGGCAGAGCCGCGC	783
Db	644	TCTCAGACACAGCCAACTACCTGCTGGCAAGAAATATTTGGCCAAAGCGCGCGAGCA	703
QY	784	CCTCCGCTGCTCATCTGCTACGTAACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	843
Db	704	CGAGGCCACAGTCATCTGCTGTAACGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	763
QY	844	TCTGACGCGCAGCTGTGGGCGCGGCTGGCAGAAACGAGCCGCGAGCTGCACCAACCCGG	903
Db	764	CTTGTCTTAACCGCTGGCGCGGAGTTGGCAGAAACGTAAGGACCTGCACCAACCCAG	823
QY	904	CGCTCTCAACGGGGCGCTTCTGTGAGGGGAGAAATGTCATGACCGCAGCTCTCTCT	963
Db	824	CCCCACTCAATGAGGTGCTTCTTGGGAGGAGCAGGCTTGGCCAGAA---GACGCTTGCA	880
QY	964	CTCTGCTTGTCTGTGGAGCGGAGCTGGAGCCGCTGGAGCAAGTGGTCCGCTGTGGGC	1023
Db	881	CCACCGTGTGCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGGTCCGCTGTGGGC	940
QY	1024	TGACTGCAACCCACTGGCGGAGCGGTGAGTCTCTGACCCAGCACCCCGCAAGAGGGG	1083
Db	941	CAGAGTGTGGCACTGGCGAGCGCGAGTGGATGGCACCCGCGCCCGCAGAACGAGGCGC	1000
QY	1084	AGGAGTCCAGGGCAGTGCCTGGACACCCGCAACTGTACAGTGCCTCTGTGTACACA	1143
Db	1001	GTGACTGACGGGACGCTACTTGACTTCAAGAACTGACCGGATGGGCTGTGGCTGTGA	1060
QY	1144	GTGCTTCTGGCC-----CTGAGGACCTGGGCC	1170
Db	1061	ATCAGAGAACTCTTAAACGACCCCTTAAAGCCGCCCTGGAGCGCTCGGAGACGTGGGC	1120
QY	1171	TCTATGTGGGCTC---ATCGCGTGGCGCTGTGCTGGTCTGCTGCTGCTGCTGCTGCTCA	1227
Db	1121	TGATGCGGCGCTGCTGCTGGCGCTCTTGTGGTCTTGGCAGTCTCTATGCTGTAGGAG	1180
QY	1228	TCTCTGTTTATTCGGGAAAGAGGGGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG	1286
Db	1181	TGATCGTGTACCGGAGAACTGCCGGAATTCGACACGGAATCACATGACTCTCTCTGTG	1240
QY	1287	--CTCAGCTCAGGCTTCCAGCGCTCAGCATCAAGCCCGCAGCAAGCAACCCCAATC	1344
Db	1241	CCCTCAGTGTGGTTTCAACCCCGTCACTTCAAGATGCAAGCCCGCAGCAACCCAGCAG	1300
QY	1345	TGCT-----CACCATCCAGCCGACCTTCAGCAACA---CCACCCTTACAGGGCAGTC	1395
Db	1301	TCTTGACCCATCCGCGCTCCGACCTTACCGGCGAGTGTGGCATCTTACCGCGAGCCTG	1360
QY	1396	TCTGTCCCGGACGA-----	1411
Db	1361	TGTATGCCCTCAGGACTCTGCCGAAGATCCCTATGACTAATTCACCCCTTCTGGATC	1420
QY	1412	-TGGGCCAGCCCAAGTTCAGCTTCAACATGGGCACTGCTGCTAGCCCC-----CC	1461
Db	1421	CCTTGCCCGAGCTCAAGATCAAGTCTATGACTCCAGCACCATCGGCTCTGGGGCTGGCC	1480
QY	1462	TGGGTGGCGCGCCGACACACTGACACAGCTCTCCCACTCTGAGGCCGAGGATTCG	1521

1481 TGGCTGATGGAGCCGACCTGCTGGGTGTTTACCAACCCGGTACATACCCAGCGGATTCTT 1540  
1522 TCTCCCGCTCTCCACCCAGAAC-----TACTTCGGT 1554  
1541 CCCGGGACACCACTTCTGACCTGCGAGCGCCAGCCCTTGGTCCACGACCTCTCTGG 1600  
1555 CCCTGCCCGGAGGACACGAGCAATGACCTATGGACCTTCAACTTCTCGGGGCGCCGCG 1614  
1601 GCCTCCCTCGAGACCCAGCAGCAGTGTGAGTGGCACCTTTGGTTGCCCTGGGTGGAGGC 1660  
1615 TGATGATCCCTAATAFACGATATCAGCTTCTCATCCCGCCAGATGCCATAACCCAGGGA 1674  
1661 TGACCATTCGCGGACACAGGGGTGAGCTTGTGTTACCAAAATGGAGCCATTCGCCAGGGA 1720  
1675 AGATCTATGAGATCTACCTCAGCTGACAGCGGAGAGAGCTGAGGTGGCCCTAGCTG 1734  
1721 AGTTCTATGACTTGTATCTAGTATCAACAGACATGAAGCACCCCTCCCATTTGGGAAG 1780  
1735 GCTGTGAGACCTGCTGAGTCCCATGTTAGTGTGGAACCCCTGGCGCTCTGCTCACCC 1794  
1781 GTTCCAGACAGTATGAGCCCTCGGTGACCTGCGGSCCAGCGGCTCTCTGTGCCC 1840  
1795 GGCGAGTACCTCGCTATGACCACTGTGGGAGCCCGAGCCCTGACAGCTGAGACCTGC 1854  
1841 GCCCTGTTGCTCAGTGTGCCCCACTGTGCTGAAGTCAATGCGGAGACTGGATCTTCC 1900  
1855 GCCTCAAAAGACGTCGTGGAGGCGAGCTGGAGCAGGAGTGTCTGCACCTGGGCGAGG 1914  
1901 AGCTCAAGACCCAGGCCCATCAGGGCCACTGGG---AGGAGTGGTGACTTTGGATGAGG 1957  
1915 AGGGCCCTCGACCTACTACTGTCAGCTGGAGGCGAGTGCCTGCTAGCTTTTACCG 1974  
1958 AGACTGTGAACACCCCTGCTACTGCGAGCTAGAGGCTAAATCTGCGCACATCTGTGG 2017  
1975 AGCAGCTGGCGCGCTTGGCCCTGGTGGAGAGGCCCTCAGGGTGGCTGCGCGCAAGCGCC 2034  
2018 ACCAGCTGGGTACTAGTGTTCACGGGCGAGTCTACTCCCGCTCGCAGTCAAGCGGC 2077  
2035 TCAAGCTGCTTCTGTTGCGCCGGTGGCTGCGACCTCCCTCGAGTACAAATCGGGTCT 2094  
2078 TCCAGCTAGCACTTCGCGCCAGCCCTCTGCACCTCCCTCGAGTATAGTCTCAGGGTCT 2137  
2095 ACTGCTGATGACACCCAGATGCACTCAAGAGGTGTGCGAGCTGGAGAGCAGCTGG 2154  
2138 ACTGTGAGGACACTCTCTGACGACTGAAGAGGTCTTAGAGCTGGAGAGCTCTGG 2197  
2155 GGGGACAGCTGATCCAGGAGCCAGGGTCTGCACTTCAAGGACAGTTACCAACCTGC 2214  
2198 GTGGCTACTTGTGGAGGAGCCAGACTTGTCTTTAAGGACAGTTACCAACCTAC 2257  
2215 GCCTATCCATCCACGATGTGCCAGCTCCCTGTGGAAGAGTGAAGTCTGTGACGTAC 2274  
2258 G-CTCTCCCTCCATGACATCCCCCATGCCCACTGGAGGAGAACTACTGGCCAGTACC 2316  
2275 AGGAGATCCCTTTATACATCTGGAATGGACGAGCGGTACTTGCACCTGCACCTTCA 2334  
2317 AGGAGATTCCTTCTACCATGTGTGAAGCGGAGCCAGAAAGCCCTGCACTGCATTCA 2376  
2335 CCCTGGAGGCTCAGCCCGACACTAGTGAAGCTGCGCTGCGAGTGTGGGTGTGGCAGG 2394  
2377 CCCTGGAGGACATAGCCCTAGCTCCACTGAGTTCACTGTAAAGTCTCGTGGCGCAGG 2436  
2395 TGGAGGGGAGCGGAGAGCTTTGAGCATCACTTCAACATCAC---CAGGACACAGGT 2451  
2437 TAGAAGGGGAGGCCAGATTTTCCAGCTGCACACACGCTGCTGAGAGCCTGTGGCT 2496  
2452 TTGCTGAGCTGTGGCTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGCCCCAGTG 2511  
2497 CCCTGGATGCACTCTGCTCTGCCCCCTGGCAATGTCTGCCACACACAGCTGGACCTATG 2556  
2512 CTTTCAAGATCCCTTCTCTATTCGGCAGAGATTAATTTCCAGCTGGAACCCACCTGTA 2571

Db 2557 CTTTCAAGATACCACTGTCCATCCCGCAAGAAAGATCTGCAACAGCCTGGAACGCCCAACT 2616  
QY 2572 GCGGGGTGCGGACTTGGGGGACTCTGGCCCAAGAAACTCCACCTGGACAGCCATCTCAGCT 2631  
Db 2617 CACGGGGCAATGACTTGGCGGCTGTTGGCAGACAGAGCTCTCCATGAGCCGGTACCTGAAT 2676  
QY 2632 TCTTTGCTTCCAGCCAGCCAGCCCAAGCCATGATCCTCAACCTGTGGAGGCGCGCACT 2691  
Db 2677 ACTTCGCCACCAAAAGTAGTCCCAAGCGGTGATCTTAGACCTCTGGGAAGCTCGGCAGC 2736  
QY 2692 TCCCCAAGCGCAACTCAGCCAGCTGGCTGAGAGTGGCTGGAGTGGGCGCAGCAGAG 2751  
Db 2737 AGGATGATGGGACCTCAACAGCCTGGCCAGTGGCTTGGAGGAGATGGGCAAGTGTGA 2796  
QY 2752 CTGGCTCTTCCACAGTGTGGAGGCTGAGTGTCTGA 2786  
Db 2797 TGCTGTAGCCATGACCACCTGATGGGATTTCTGA 2831

## RESULT 7

US-09-306-902A-3  
; Sequence 3, Application US/09306902A  
; Patent No. 6277585  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsey  
; Masu, Masayuki  
; Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/306,902A  
; FILING DATE: 07-May-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2831 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-306-902A-3

Query Match 27.2%; Score 783.4; DB 3; Length 2831;  
Best Local Similarity 59.4%; Pred. No. 5.7e-161;  
Matches 1625; Conservative 0; Mismatches 971; Indels 139; Gaps 12;  
QY 184 ACCAGTGCCTGTGCGCAACCGGAGCCTGCTTCCCACTTCTGTGTGGAGCCGAGGATG 243  
Db 104 ACTCTTCCCATCAGACCCCGGAGAGCTGCTCACTTCTGTGGAACGAGAGATG 163  
QY 244 TGTATATGTCAGAGAACAGCAGTGTGTGTGTCAGAGCGCGTGTGTCGCCCGCAGCAGA 303



QY	2335	CCCTGGAGCGTGTGACGCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGG	2394
Db	2377	CCCTGGAGAGACATAGCCCTAGCCTCCACTGAGTTTACCTGTAAGGTCTGCGTGGCGCAGG	2436
QY	2395	TGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCAC---CAAGACACAAGGT	2451
Db	2437	TAGAAGGGGAAGGCCAGAGATTTTCAGCTGCACACCAACGCTGGCTGAGACGCCTCTCGCT	2496
QY	2452	TTGCTGAGCTGCTGGCTCTGAGAGTGAAGCGGGGTCCACAGCCCTGGTGGGCCCCACGTG	2511
Db	2497	CCCTGGATGCACTCTGCTCTGCCCTGGCAATGCTGCCACACACAGCTGGGACCCCTATG	2556
QY	2512	CTTTCAAGATCCCTCTCTCATTCGGCAGAAGATAATTTCCAGCCTGGAACCCACCTGTA	2571
Db	2557	CTTTCAAGATACCACTGTCCATCCGCAGAGAGATCTGCAACAGCTGGAGCGCCCAACT	2616
QY	2572	GGCGGGTCCGACTGGCGGACTCTGGCCAGAAACTCCACCTGGACAGACCATCTCAGCT	2631
Db	2617	CACGGGCAATGACTGGCGGCTGTGGCACAAGACTTCCATGGACCGGTACTTGAAC	2676
QY	2632	TCATTGCTCCAGCCAGCCACAGCATGATCTCAACCTGTGGAGGCGCGCACT	2691
Db	2677	ACTTCGGCCACCAAGCTAGTCCACAGGCGTGATCTTAGACCTCTGGGAAGCTCGGCAGC	2736
QY	2692	TCCCAACGGCAACTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCACGCCAGCG	2751
Db	2737	AGGATGATGGGACCTCAACAGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGTGA	2796
QY	2752	CTGSCCTCTTCACAGTCTCGAGGCTGAGTGCTGA	2786
Db	2797	TGCTGGTAGCCATGACCACCTGATGGCGATTGCTGA	2831

## RESULT 8

```

US-09-833-381-1807
; Sequence 1807, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1807
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1605)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1807

```

	Query Match	14.5%;	Score 419;	DB 4;	Length 1605;
	Best Local Similarity	62.4%;	Prod. No. 6.6e-82;		
	Matches 737;	Conservative	0;	Mismatches 435;	Indels 10; Gaps 5;
QY	1556	CCTGCCCGGAGGCACCAACATGACCTATGGAGCCTTCAACTCTCGGGGGCGGGCT	1615		
Db	421	CCTGCCCGGAGACCCAGGGAGCAGGCTCAGCGGCACCTTTGGTGCTGGTGGGAGGCT	480		
QY	1616	GATGATCCCTAATACAGGTATCAGCTCTCTCATCCCCCAGATGCCATACCCCGAGGAA	1675		
Db	481	CANCATCCCCGGGCACAGGGGTCAAGTTGTGGTGCCTCAATGAGCCATATCCCGAGGC	540		
QY	1676	GATCTATGA--GATCTACCTCAGCTGCACAGCCGGAACGCTGAGTTGCCCTTAGCT	1733		
Db	541	GTCTACGAGATGTTATTCTATCATCAACAGGCAGAAAGTACCTCCCGCTTTCAGAA	600		

QY	1734	GGCTGTCAGACCCCTGCTGAGTGCCCATCTGATTAGCTGTGTGAGACCCCTGGCGTCTCTGTCACC	1793
Db	601	GGGACCCACACAGTATTGAGCCCTCGGTGACCTGTGGACCCACAGGCTCTCTGCTGTGC	660
QY	1794	CGGCCAGTCATCCTGGCTATGGAACAATGTGGGAGGCCACGCCCTGACAGCTGAGCGCTG	1853
Db	661	CGCCCCGTATCCTCACCATGCCCCACTGTGCCGAAGTCAGTGGCCCGTGACTGCGATCTTT	720
QY	1854	CGCCTCAAAAACAGTCGTGCGAGGGCAGCTGGCAGCAGGATGTGTCGACCTCTGGGGCGAG	1913
Db	721	CAGCTCAAGACCCAGGCCCCACAGGGCCACTGG--ANGAGGTGGTGAACCTTGGATGAG	777
QY	1914	GAGGGCCCTCCCACTCTACTATGCCAGCTGGAGGCCAGTGCGCTGCTACGTCTTCACC	1973
Db	778	GAGACCTTGAACACACACCTGCTACTGCGAGCTGGAGCCCAAGGCCCTGTGCATCTCTGCTG	837
QY	1974	GAGCAGCTGGGCGCGT-TTGGCCCTGFTGGAGAGGCCCTCAGCGTGCCTGGCGCCCAAGCG	2032
Db	838	GACACAGCTGGGCACTACCGTGTTCACGGGCGAGTCTTATTCGCCGCTCAGCAGTCAAGCG	897
QY	2033	CCTCAAGCTGCTTCTGTTTGGCGCGGTGGCGTGCACCTCCCTCGAGTACAAACATCCGGGT	2092
Db	898	GCTCCAGCTGGCGCTTCTTGGCCCGCCCTCTGCACTCCCTGGAGTACAGCTCTCGGGT	957
QY	2093	CTACTGCTGCATCACACCCACAGTACGACTCAGGAGGTGGTGAGCTGGAGAGCAGCT	2152
Db	958	CTACTGCTGGAGACACGGCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGCACTCT	1017
QY	2153	GGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGAGCAGTTACCAACACCT	2212
Db	1018	GGGCGGATCTTGGTGGAGAGCGGAACCGCTTAATGTTCAAGGACAGTTACCAACCT	1077
QY	2213	GGCGCTATCCATCCAGATGCGCAGTCCCTGTGGAGAGTAAGCTCTTGTCAAGCTA	2272
Db	1078	GGCGCTCTCCCTCCATGACCTCGCCCATGCCATTGGAGGAGCAAGCTGTGSCCAATA	1137
QY	2273	CCAGGAGATCCCTTTTATCACATCTGGATGGACGCGACGCGGTACTTGCATGCACTT	2332
Db	1138	CCAGGAGATCCCTTTTATCACATTTGGAGTGGCAGCCAGAGGCCCTCCATGCACTTT	1197
QY	2333	CACCTGGAGCGTGTGAGCCCCAG-CACTAGTGACCTGGCTGCAAGCTGTGGGTGCGC	2391
Db	1198	CACCTGGAGAGCACAAGCTTGGCTTCCACAGAGCTCACCTGCAAGATCTGGGTGCGC	1257
QY	2392	AGGTGAGGCGAGCGGACAGCTTCAGCATCACTTCAACATCAACAG--GACACAA	2448
Db	1258	AAGTGAAGGGGAGGCGCAGATATTCAGCTGCATACCACTCTGGCAGAGACACCTGCTG	1317
QY	2449	GGTTTGTCAGCTGCTGGCTCTGGAGAGTGAAGCGGGGTGCCAGCCCTGGTGGCCCCA	2508
Db	1318	GCTCCCTGGACATCTCTGCTCTGCCCCCTGGCAGCACTGTCAACCCAGCTGGGACCTT	1377
QY	2509	GTGCTTCAAGATCCCTTCTCATTPCGCAGAAGATAATTTCCAGCCTGGAGCCACCT	2568
Db	1378	ATGCTTCAAGATCCCACTGTGCATCCGCGCAGAAGATATGCAACAGCTAGATGCCCA	1437
QY	2569	GTAGGCGGGTGGCGACTGGCGGACTCTGGCCCGAGAACTCCACTGGACAGCCATCTCA	2628
Db	1438	ACTCACGGGGCAATGACTGGCGGATGTAGCACAGAAGCTCTATGACCCGGTACCTGA	1497
QY	2629	GCTTTTTCCTCCAGGCCACGCCCAAGCCATGATCTCAACCTGTGGAGCGCGGC	2688
Db	1498	ATTACTTTGCCACCAAGGAGGCCCAAGGNTGTGATCTTGACCTCTGGAGAGCTGTGC	1557
QY	2689	ACTTCCCACAGCAGCTCAGCCAGCTGGCTGACAGAGTG	2730
Db	1558	AGCAGACGATGGGACCTCAACAGCTGCGAGTGCCTTG	1599

RESULT 9  
US-08-253-155A-17/c  
; Sequence 17, Application US/08253155A



Patent No.	5691147
GENERAL INFORMATION:	
APPLICANT:	Gyuris, Jeno
APPLICANT:	Draetta, Giulio
TITLE OF INVENTION:	CDK4 Binding Proteins
NUMBER OF SEQUENCES:	95
CORRESPONDENCE ADDRESS:	
ADDRESSEE:	LAHIVE & COCKFIELD
STREET:	60 State Street
CITY:	Boston
STATE:	MA
COUNTRY:	USA
ZIP:	02109
COMPUTER READABLE FORM:	
MEDIUM TYPE:	Floppy disk
COMPUTER:	IBM PC compatible
OPERATING SYSTEM:	PC-DOS/MS-DOS
SOFTWARE:	ASCII(text)
CURRENT APPLICATION DATA:	
APPLICATION NUMBER:	US/08/253,155A
FILING DATE:	02-JUN-1994
CLASSIFICATION:	435
ATTORNEY/AGENT INFORMATION:	
NAME:	Vincent, Matthew P.
REGISTRATION NUMBER:	36,709
REFERENCE/DOCKET NUMBER:	MII-028
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	(617) 227-7400
TELEFAX:	(617) 227-5941
INFORMATION FOR SEQ ID NO:	17:
SEQUENCE CHARACTERISTICS:	
LENGTH:	771 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	single
TOPOLOGY:	linear
MOLECULE TYPE:	cdna
US-08-253-155A-17	
Query Match	8.9%; Score 256; DB 1; Length 771;
Best Local Similarity	67.2%; Pred. No. 1.3e-46;
Matches	406; Conservative 1; Mismatches 191; Indels 6; Gaps 3;
QY	1741 AGACCCGTGAGTCCCATCGTTAGCTGTGGAGCCCCCTGGCTCCTCACCCGCCAG 1800
DB	738 AGACAGTAATTGAGCCCTCGGTGACCTGTGGAGCCACAGCCCTCTGTGTCGCCGCCG 679
QY	1801 TCATCTGGCTATGACACACTGTGGGAGGCCACCCCTGACAGCTGGAGCCTGCCTCA 1860
DB	678 TCATCTCACCATGCCCACTGTGCCGANGTAGTGCCGTGACTGGATCTTTTACGTCA 619
QY	1861 AAAAGCAGTGTGGCAGGGAGCTGGAGCAGAGATGTGTGACCTGGCGGAGGGCGC 1920
DB	618 AGACCCAGGCCACCCAGGGCCACTTGGG---AGGAGTGGTGACCTGGATGAGGAGCCC 562
QY	1921 CCTCCACCTCTACTGTCAGCTGAGGCCAGCTGCTGTACGCTTTTACCGAGCAGC 1980
DB	561 TGAACACACCTGTCTATGACAGCTGGAGCCAGGCGCTGTACATCCTGTGGACCAGC 502
QY	1981 TGGGCGCGTTTTGCCCTGGTGGAGAGCCCTCAGCGTGGCTGGCGGCAAGCGCTCAAGC 2040
DB	501 TGGGCACCTACGTGTTCAGGGCGAGTCTTATTCGCGCTCAGCAGTCAAGCGGCTCCAGC 442
QY	2041 TGCTTCTGTTTTGCCCGGTGSCCTGCACTTCCCTCGAGTACAACATCCGGGTCTACTGCC 2100
DB	441 TGGCGGTCTTTCGCGCGCGCCCTCTGCACTTCCCTGGAGTACAGCTTCCGGGTCTACTGTC 382
QY	2101 TGCATGACACCCAGCATCACTCAAGGAGTGTGTGACCTGGAGAGAGCAGTGGGGGAC 2160
DB	381 TGGAGGACAGGCGCTGTAGCACTGAAGGAGGTGTGTGGAGCTGGAGCGGACTCTGGCGGAT 322
QY	2161 AGCTGATCAGAGGACCGGGTCTGCACTTCAAGGACAGTTACCAACCTTCGCGCTAT 2220
DB	321 ACTTGGTGGAGGCGCGAAACCGCTTAATGTTCAGGACAGTTAACACACTTCGCG-CTCT 263

	Patent No. 5691147	
	GENERAL INFORMATION:	
	APPLICANT: Gyuris, Jeno	
	APPLICANT: Draetta, Giulio	
	TITLE OF INVENTION: CDK4 Binding Proteins	
	NUMBER OF SEQUENCES: 95	
	CORRESPONDENCE ADDRESS:	
	ADDRESSEE: LAHIVE & COCKFIELD	
	STREET: 60 State Street	
	CITY: Boston	
	STATE: MA	
	COUNTRY: USA	
	ZIP: 02109	
	COMPUTER READABLE FORM:	
	MEDIUM TYPE: Floppy disk	
	COMPUTER: IBM PC compatible	
	OPERATING SYSTEM: PC-DOS/MS-DOS	
	SOFTWARE: ASCII(text)	
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER: US/08/253,155A	
	FILING DATE: 02-JUN-1994	
	CLASSIFICATION: 435	
	ATTORNEY/AGENT INFORMATION:	
	NAME: Vincent, Matthew P.	
	REGISTRATION NUMBER: 36,709	
	REFERENCE/DOCKET NUMBER: MII-028	
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE: (617) 227-7400	
	TELEFAX: (617) 227-5941	
	INFORMATION FOR SEQ ID NO: 17:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 771 base pairs	
	TYPE: nucleic acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	MOLECULE TYPE: cDNA	
	US-08-253-155A-17	
	Query Match 8.9%; Score 256; DB 1; Length 771;	
	Best Local Similarity 67.2%; Pred. No. 1.3e-46;	
	Matches 406; Conservative 1; Mismatches 191; Indels 6; Gaps 3;	
QY	1741 AGACCCGTGAGTCCCCTGGTGGAGGCCCGCCTGCCTCACCGGCCAG 1800	
DB	738 AGACAGTAATGAGCCCCCTGGTGACCTGTGGACCCACAGCCCTCTGTGTCGCCGCCCG 679	
QY	1801 TCATCTGGCTATGACACACTGTGGGAGGCCCGCCTGACAGCTGGAGCCTGCCTCA 1860	
DB	678 TCATCTCACCATGCCCACTGTGCCGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619	
QY	1861 AAAAGCAGTGTGGAGGGAGCTGGAGCAGATGTGTGACCTGGCGGAGGGCGC 1920	
DB	618 AGACCCAGGCCCAACAGGGCCACTTGGG---AGGAGTGGTGACCTGGATGAGGAGCCC 562	
QY	1921 CCTCCACCTCTACTGTCAGCTGGAGGCCAGTCTGCTACGCTTTTCCAGGACGC 1980	
DB	561 TGAACAACCTCTGCTGTCAGCTGGAGGCCCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502	
QY	1981 TGGGCGCGCTTTGCCCTGGTGGAGAGCCCTCAGCGTGGCTGGCGGCAAGCGCCTCAAGC 2040	
DB	501 TGGGCACCTACGTGTTCAGGGCGAGTCTTATTCGCGCTCAGCAGTCAAGCGGCTCCAGC 442	
QY	2041 TGCTTCTGTTTTGCCCGGTGCGCTGCACTCCCTCCGAGTACAACTCCGGGTCTACTGCC 2100	
DB	441 TGGCGGTCTTGGCGCGCGCCCTCTGCACTCTCTGGAGTAGACACTTCCGGGTCTACTGTC 382	
QY	2101 TGCATGACACCAAGATCACTCAAGGAGTGTGTGACCTGGAGAGAGCAGTGGGGGAC 2160	
DB	381 TGGAGGACAGGCTGTAGCACTGAAGGAGGTGTGGAGCTGGAGCGGACTCTGGGCGGAT 322	
QY	2161 AGCTGATCAGAGGACCGGTCTGCACTTCAAGGACAGTTCACCACTCGCGCTAT 2220	
DB	321 ACTTGGTGGAGGCGCGAAAACCGCTTAATGTTCAGGACAGTTAACCACTGCG-CTCT 263	



QY 2086 TCCGGGTCTACTGCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGA 2145  
 DB 178 TCCGGGTCTACTGCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGA 2145  
 QY 2146 AGCAGCTGGGGGAGAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGTTACC 2205  
 DB 238 GACTCTGGGGGAGATCTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACC 2297  
 QY 2206 ACAACCT 2212  
 DB 298 ACAACCT 304

RESULT 11  
 US-09-306-902A-4  
 ; Sequence 4, Application US/09306902A  
 ; Patent No. 6277585  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tessier-Lavigne, Marc  
 ; Leonardo, E. David  
 ; Hink, Lindsay  
 ; Masu, Masayuki  
 ; Kazuko, Keino-Masu  
 ; TITLE OF INVENTION: Netrin Receptors  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 268 BUSH STREET, SUITE 3200  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/306,902A  
 ; FILING DATE: 07-May-1999  
 ; CLASSIFICATION: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OSMAN, RICHARD A  
 ; REGISTRATION NUMBER: 36,627  
 ; REFERENCE/DOCKET NUMBER: UC96-217  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 343-4341  
 ; TELEFAX: (415) 343-4342  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 305 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 US-09-306-902A-4

Query Match 4.1%; Score 119; DB 3; Length 305;  
 Best Local Similarity 68.1%; Pred. No. 5.6e-17;  
 Matches 209; Conservative 0; Mismatches 95; Indels 3; Gaps 3;  
 QY 1906 TGGGCGAGGAGGCGCCCTCCACCTCTACTGCGAGCTGGAGGCGAGTGGCTGTAGC 1965  
 DB 1 TGGATGAGAGACCTTGACACACCTGCTACTG-CAGCTGGAGCCAGGCGCTG-TACA 58  
 QY 1966 TCTTACCGAGAGCTGGCGCGCTTTGCCCTGTGGGAGAGGCGCTTACGCGTGGCTGCCG 2025  
 DB 59 TCTTGTGACAGCTGGGCGACCTAGCTTTTACGGGCGAGTCTTATTCCTCCCTAGCAG 118  
 QY 2026 CAAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGCTGCACTCCCTCTCGAGTACAACA 2085

DB 119 TCAAGGGCTCCAGCTGGCCGT-TTCGCCCGCGCCCTCTGCACCTCCCTGGAGTACAGCC 177  
 QY 2086 TCCGGGTCTACTGCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGA 2145  
 DB 178 TCCGGGTCTACTGCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGA 2145  
 QY 2146 AGCAGCTGGGGGAGAGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGTTACC 2205  
 DB 238 GACTCTGGGGGAGATCTTGGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTACC 2297  
 QY 2206 ACAACCT 2212  
 DB 298 ACAACCT 304

RESULT 12  
 US-08-483-533-38  
 ; Sequence 38, Application US/08483533  
 ; Patent No. 6172047  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Roizman, Bernard  
 ; APPLICANT: Chou, Joany  
 ; TITLE OF INVENTION: Method for Treating Tumorigenic  
 ; TITLE OF INVENTION: Diseases  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/483,533  
 ; FILING DATE: 07-MAR-95  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/419,853  
 ; FILING DATE: 11-APR-95  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/861,233  
 ; FILING DATE: 31-MAR-92  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, James P.  
 ; REGISTRATION NUMBER: 28,491  
 ; REFERENCE/DOCKET NUMBER: 28097/32742  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312/474-6300  
 ; TELEFAX: 312/474-0448  
 ; TELEX: 25-3856  
 ; INFORMATION FOR SEQ ID NO: 38:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1280 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-08-483-533-38

Query Match 1.8%; Score 51.8; DB 3; Length 1280;  
 Best Local Similarity 51.0%; Pred. No. 0.034;  
 Matches 122; Conservative 0; Mismatches 117; Indels 0; Gaps 0;  
 QY 6 GGGCTCCGGGTGAGCGCTAAAGCGGCTCCCGCGCGGCGCCCGCGCGCGCGCGCG 65  
 DB 555 GCGGCGGAGGCGCGCGCGAGCCCGCGGACCCCTCGGACCCCGCGGACCCCGCG 614  
 QY 66 CCCGCTGCCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 125





**This Page Blank (uspto)**

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	2881	100.0		2881	11	US-09-970-944-1	Sequence 1, Appli
2	2676.4	92.9		2752	13	US-09-918-779-1	Sequence 1, Appli
3	2676.4	92.9		2752	17	US-10-824-932-1	Sequence 1, Appli
4	2402.6	83.4		3580	17	US-10-911-623-13	Sequence 13, Appli
5	2235.8	77.6		3014	10	US-09-933-261-1	Sequence 1, Appli
6	2235.8	77.6		3014	15	US-10-256-702-1	Sequence 1, Appli
7	2189.4	76.0		2697	15	US-10-340-154-15	Sequence 15, Appli
8	1570.4	54.5		1787	10	US-09-933-261-2	Sequence 2, Appli
9	1570.4	54.5		1787	15	US-10-256-702-2	Sequence 2, Appli
10	1193.6	41.4		1321	13	US-10-296-115-365	Sequence 365, App
11	1147.8	39.8		1282	9	US-09-833-381-1806	Sequence 1806, Ap
12	889	30.9		2860	13	US-10-087-684-1	Sequence 1, Appli
13	889	30.9		2860	13	US-10-918-779-1	Sequence 1, Appli
14	887.4	30.8		2860	13	US-10-087-684-3	Sequence 3, Appli

180

Db 121 TGGGCATAGTCTCGCCGCTTGGCTCCGCGGCTCGGGTCCCGACAGAGTGCCACCGTGG 180  
Qy 181 CCAACCCAGTGTCTGTGTCACACCCCGAGACCTGTCTCCCACTTCTCTGGTGGAGCCCGAGG 240  
Db 181 CCAACCCAGTGTCTGTGTCACACCCCGAGACCTGTCTCCCACTTCTCTGGTGGAGCCCGAGG 240  
Qy 241 ATGTGTACATCGTCAAGAACAGACCAAGTGTCTGTGTGTCAGAGCGGTGCGCCGACGC 300  
Db 241 ATGTGTACATCGTCAAGAACAGACCAAGTGTCTGTGTGTCAGAGCGGTGCGCCGACGC 300  
Qy 301 AGATCTTCTCAAGTGTCAAGCGGAGTGGTGGCGCAGGTGGACCACTGATCGAGCGCA 360  
Db 301 AGATCTTCTCAAGTGTCAAGCGGAGTGGTGGCGCAGGTGGACCACTGATCGAGCGCA 360  
Qy 361 GCACAGACGGAGTGTGTGAGCCGACCACTGAGGTTCGCAATTAATGTCTCAAGGCGC 420  
Db 361 GCACAGACGGAGTGTGTGAGCCGACCACTGAGGTTCGCAATTAATGTCTCAAGGCGC 420  
Qy 421 AGGTGAGAAAGTGTTCGGGCTGGAGGAATACTGGTCCAGTGGGATGAGGTCTCT 480  
Db 421 AGGTGAGAAAGTGTTCGGGCTGGAGGAATACTGGTCCAGTGGGATGAGGTCTCT 480  
Qy 481 CGGGCACCAACAGAGTCAGAGGCTTACATCCGATAGCAGATTCGCGCAAGAACTTCG 540  
Db 481 CGGGCACCAACAGAGTCAGAGGCTTACATCCGATAGCAGATTCGCGCAAGAACTTCG 540  
Qy 541 AGCAGAGCGCTGGCCCAAGGAGTGTCCCTGGAGCAGGGCATCGTCTGCCCTGCCGTC 600  
Db 541 AGCAGAGCGCTGGCCCAAGGAGTGTCCCTGGAGCAGGGCATCGTCTGCCCTGCCGTC 600  
Qy 601 CACCGAGGGCATCCCTCAGCCAGGAGTGGCTCCGGAACGAGGACCTGGTGACC 660  
Db 601 CACCGAGGGCATCCCTCAGCCAGGAGTGGCTCCGGAACGAGGACCTGGTGACC 660  
Qy 661 CGTCCCTGGACCCCAATGTATACATACCGCGGAGCAGAGCTGGTGGAGAGGCGCC 720  
Db 661 CGTCCCTGGACCCCAATGTATACATACCGCGGAGCAGAGCTGGTGGAGAGGCGCC 720  
Qy 721 GCCTTGTGTGACACGGCCAACTACACCTCGTGGCCCAAGAAATCGTGCGCACGTGGCGCA 780  
Db 721 GCCTTGTGTGACACGGCCAACTACACCTCGTGGCCCAAGAAATCGTGCGCACGTGGCGCA 780  
Qy 781 GCGCTCTCGTGTCTGTCTACGTGAACTGAGTGGTGGTGGAGTGGTGGAGTGGT 840  
Db 781 GCGCTCTCGTGTCTGTCTACGTGAACTGAGTGGTGGTGGAGTGGTGGAGTGGT 840  
Qy 841 CCGTCTGACGCGCCAGCTGTGGCGCGCTGGCAGAAACGGAGCGGAGCTGCCAACCC 900  
Db 841 CCGTCTGACGCGCCAGCTGTGGCGCGCTGGCAGAAACGGAGCGGAGCTGCCAACCC 900  
Qy 901 CGCGCGCTCTCAACCGGGCGCTTTCTGTGAGGGCGAGATGTCCATGACCGCACCGTCT 960  
Db 901 CGCGCGCTCTCAACCGGGCGCTTTCTGTGAGGGCGAGATGTCCATGACCGCACCGTCT 960  
Qy 961 CTTCTCTGCTGTCTGTGTGAGCGAGCTGGAGCCGCTGGAGCAAGTGTGGCGCTGTG 1020  
Db 961 CTTCTCTGCTGTCTGTGTGAGCGAGCTGGAGCCGCTGGAGCAAGTGTGGCGCTGTG 1020  
Qy 1021 GGTCTGAGTGCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCAGCCCGCAACGGAG 1080  
Db 1021 GGTCTGAGTGCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCAGCCCGCAACGGAG 1080  
Qy 1081 GGGAGGAGTGCACGGGCACTGACCTGGACACCGCAACTGTGTACAGTGTGATGTGTGAT 1140  
Db 1081 GGGAGGAGTGCACGGGCACTGACCTGGACACCGCAACTGTGTACAGTGTGATGTGTGAT 1140  
Qy 1141 AAGTCTCTTGGCCCTGAGGAGTGGCCCTTATGTGGGCTCTATCGCGGTGGCGGTCT 1200  
Db 1141 AAGTCTCTTGGCCCTGAGGAGTGGCCCTTATGTGGGCTCTATGTGGGCTCTATCGCGGTGGCGGTCT 1200  
Qy 1201 GCTGTGTCTGTGTCTGTGTCTCTCATCTGTCTTATTTGCCGGAAGAGGAGGCGTGG 1260  
Db 1201 GCTGTGTCTGTGTCTGTGTCTCTCATCTGTCTTATTTGCCGGAAGAGGAGGCGTGG 1260

Qy 1261 ACTCAGATGTGTGCTACTGTCTCACTTCTCACTCAGGCTTCCAGCCCGTCAAGATCAAGC 1320  
Db 1261 ACTCAGATGTGTGCTACTGTCTCACTTCTCACTCAGGCTTCCAGCCCGTCAAGATCAAGC 1320  
Qy 1321 CCAGCAAGCAGACAAACCCCATCTGTCTCACTCAGGCGGACCTCAGACCAACCA 1380  
Db 1321 CCAGCAAGCAGACAAACCCCATCTGTCTCACTCAGGCGGACCTCAGACCAACCA 1380  
Qy 1381 CTTACAGGGCAGTCTGTCTCCCGCAGGATGGGCGCCAGCCCAAGTTCACGCTCACCA 1440  
Db 1381 CTTACAGGGCAGTCTGTCTCCCGCAGGATGGGCGCCAGCCCAAGTTCACGCTCACCA 1440  
Qy 1441 ATGGGCACTGTCTCAGCCCGCTGGGTGGGGCGCCACACACTGCAACAGCTCTCCCA 1500  
Db 1441 ATGGGCACTGTCTCAGCCCGCTGGGTGGGGCGCCACACACTGCAACAGCTCTCCCA 1500  
Qy 1501 CTTCTGAGGCGGAGGAGTGTCTCTCCCGCTCTCCAGCCAGAACTACTTCCGCTTCCCTGC 1560  
Db 1501 CTTCTGAGGCGGAGGAGTGTCTCTCCCGCTCTCCAGCCAGAACTACTTCCGCTTCCCTGC 1560  
Qy 1561 CCGAGGCAACAGCACTGACCTATGGGACCTTCACTTCCCTCGGGGCGCGCTGATGA 1620  
Db 1561 CCGAGGCAACAGCACTGACCTATGGGACCTTCACTTCCCTCGGGGCGCGCTGATGA 1620  
Qy 1621 TCCCTAATACAGGTATCAGGCTCTCTCATCCCGCCAGATGCCATAACCCGAGGGAAGTCT 1680  
Db 1621 TCCCTAATACAGGTATCAGGCTCTCTCATCCCGCCAGATGCCATAACCCGAGGGAAGTCT 1680  
Qy 1681 ATGAGATCTACTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 1740  
Db 1681 ATGAGATCTACTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 1740  
Qy 1741 AGACCTCTGAGTCCCATCTGTAGCTGTGACCCCTGGCGTCTCTGCTCACCAGGCGAG 1800  
Db 1741 AGACCTCTGAGTCCCATCTGTAGCTGTGACCCCTGGCGTCTCTGCTCACCAGGCGAG 1800  
Qy 1801 TCATCTCTGGTATGAGACCTGTGGGAGCCAGCCCTGACAGCTGAGGCTGGGCTGCTCA 1860  
Db 1801 TCATCTCTGGTATGAGACCTGTGGGAGCCAGCCCTGACAGCTGAGGCTGGGCTGCTCA 1860  
Qy 1861 AAAAGCAGTGTGGAGGCGAGCTGGGAGGAGGATGTCTGACCTGGGCGGAGGAGGCGC 1920  
Db 1861 AAAAGCAGTGTGGAGGCGAGCTGGGAGGAGGATGTCTGACCTGGGCGGAGGAGGCGC 1920  
Qy 1921 CCTCCACCTCTACTCTGCTGAGGCGAGTGGCTGCTGCTGCTTCCAGGAGGAGC 1980  
Db 1921 CCTCCACCTCTACTCTGCTGAGGCGAGTGGCTGCTGCTGCTTCCAGGAGGAGC 1980  
Qy 1981 TGGGCGGCTTTGGCTGGTGGGAGGCGCTCAGGCTGGTGGCGCCAGGCGCTCAAGC 2040  
Db 1981 TGGGCGGCTTTGGCTGGTGGGAGGCGCTCAGGCTGGTGGCGCCAGGCGCTCAAGC 2040  
Qy 2041 TGTCTCTGTTGGCGCGTGGCTGCACTCCCTGAGTACAACTCCGGGTCTACTGCC 2100  
Db 2041 TGTCTCTGTTGGCGCGTGGCTGCACTCCCTGAGTACAACTCCGGGTCTACTGCC 2100  
Qy 2101 TGCATGACACCCAGCATCAGCTCAAGGAGTGGTGGAGGAGGAGGAGGAGGAG 2160  
Db 2101 TGCATGACACCCAGCATCAGCTCAAGGAGTGGTGGAGGAGGAGGAGGAGGAG 2160  
Qy 2161 AGCTGATCCAGGAGCAGCGGTCTTGCACCTTCAAGGAGTGGTGGAGGAGGAGGAGGAG 2220  
Db 2161 AGCTGATCCAGGAGCAGCGGTCTTGCACCTTCAAGGAGTGGTGGAGGAGGAGGAGGAG 2220  
Qy 2221 CCATCCAGATGTGGCGAGCTCCCTGTGGAAGAGTAACTCTGTGACCTACAGGAGGAG 2280  
Db 2221 CCATCCAGATGTGGCGAGCTCCCTGTGGAAGAGTAACTCTGTGACCTACAGGAGGAG 2280  
Qy 2281 TCCCTCTTTTATCACAATCTGGAATGGCAGCAGCGGTACTTGCATGCACTTACCGCTGG 2340  
Db 2281 TCCCTCTTTTATCACAATCTGGAATGGCAGCAGCGGTACTTGCATGCACTTACCGCTGG 2340







```

; Sequence 1, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shinkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsbrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glennda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/10/624,932
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 09/918,779
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2752
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-624-932-1

Query Match          92.9%; Score 2676.4; DB 17; Length 2752;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2728; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

QY      42      CCGCGGGGGCCCCCGCGCCCGCGCGCTCTCTGGGCATAGTCTCTCGCGCTTTGGTTCGCGCGACTCGGGTGCC 101
Db      1      CCGCGGGGGCCCCCGCGCCCGCGCGCTCTCTGGGCATAGTCTCTCGCGCTTTGGTTCGCGCGACTCGGGTGCC 60

QY      102     GGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTTGGTTCGCGCGACTCGGGTGCC 161
Db      61     GGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTTGGTTCGCGCGACTCGGGTGCC 120

QY      162     CACGACAGTGCCACCGTGGCCAAACCCAGTGCCTTGGTGCCAAACCCGACCTGTCTTCCCCAC 221
Db      121     CACGACAGTGCCACCGTGGCCAAACCCAGTGCCTTGGTGCCAAACCCGACCTGTCTTCCCCAC 180

QY      222     TTCCTGTGTGAGCCCGAGGATGTGTACATCGTCAAGAAACGACGAGTGTGTGTGTGTC 281

```

QY	1362	GACCTCAG---CACACACACACCTACACGGCAGCTCTGTCCCCGGCAGGATGGGCC	1418
DB	1318	GACCTCAGACACACACACACACCTACACAGGCGAGTCTCTGTCCCGGCGAGATGGGCC	1377
QY	1419	AGCCCCAAGTTCCAGCTCACCAATGGGCACTTGCTCAGCCCCCTGGGTGGCGGCGCAC	1478
DB	1378	AGCCCCAAGTTCCAGCTCACCAATGGGCACTGCTCAGCCCCCTGGGTGGCGGCGCAC	1437
QY	1479	ACACTGCACACAGCTCTCCACCTCTGAGGGCGGAGGAGTTGCTCTCCGGCTCTCCACC	1538
DB	1438	ACACTGCACACAGCTCTCCACCTCTGAGGGCGGAGGAGTTGCTCTCCGGCTCTCCACC	1497
QY	1539	CAGAACTACTTCCGCTCCCTGCCCGGAGCACACAGCAACATGACCTATGGGACCTTCAAC	1598
DB	1498	CAGAACTACTTCCGCTCCCTGCCCGGAGGCAACAGCAACATGACCTATGGGACCTTCAAC	1557
QY	1599	TTCTCTGGGGCGCGGTGATGCCCTAATACAGGTATCAGGCTCTCTCATGCCCCAGAT	1658
DB	1558	TTCTCTGGGGCGCGGTGATGCCCTAATACAGGTATCAGGCTCTCTCATGCCCCAGAT	1617
QY	1659	GCCATACCCCGAGGAGATCTATGAGATCTACTCAGCTGCACAAAGCCGGAACGCTG	1718
DB	1618	GCCATACCCCGAGGAGATCTATGAGATCTACTCAGCTGCACAAAGCCGGAACGCTG	1677
QY	1719	AGGTTGCCCTAGCTGGCTCTCAGACCTCTGAGTCCCATCGTTAGCTGTGGACCCGCT	1778
DB	1678	AGGTTGCCCTAGCTGGCTCTCAGACCTCTGAGTCCCATCGTTAGCTGTGGACCCGCT	1737
QY	1779	GGGCTCTGCTACCCGGCGAGTCTACTGGCTATGGAACACTGTGGGAGCCGACGCT	1838
DB	1738	GGGCTCTGCTACCCGGCGAGTCTACTGGCTATGGAACACTGTGGGAGCCGACGCT	1797
QY	1839	GACAGCTGGAGCCTGCGCCTCAAAAGCAGTCTGTGGAGGCGAGCTGGGAGCAGATGTG	1898
DB	1798	GACAGCTGGAGCCTGCGCCTCAAAAGCAGTCTGTGGAGGCGAGCTGGG--AGGATGTG	1854
QY	1899	CTGCACCTGGCGAGGAGGCGCCTCCCACTCTACTATGCGAGCTGAGGCGAGTGCC	1958
DB	1855	CTGCACCTGGCGAGGAGGCGCCTCCCACTCTACTCTGCGAGCTGGAGGCGAGTGCC	1914
QY	1959	TGCTTACGCTCTTACCGAGAGCTGGGCGGCTTTGCCCTGGTGGGAGGCGCTCAGCGTG	2018
DB	1915	TGCTTACGCTCTTACCGAGAGCTGGGCGGCTTTGCCCTGGTGGGAGGCGCTCAGCGTG	1974
QY	2019	GCTGCCGCAAGCGCCTCAAGCTGCTTCTGTTTGGCGCGGTGGCTGTGCACCTCCCTCGAG	2078
DB	1975	GCTGCCGCAAGCGCCTCAAGCTGCTTCTGTTTGGCGCGGTGGCTGTGCACCTCCCTCGAG	2034
QY	2079	TACAACATCCGGGTCTACTGCTGCATGACACCCAGATGCATCAAGGAGGTGGTGCGAG	2138
DB	2035	TACAACATCCGGGTCTACTGCTGCATGACACCCAGATGCATCAAGGAGGTGGTGCGAG	2094
QY	2139	CTGGAGAGAGCTGGGGGGACAGCTGATCCAGAGGCCACGGGTCTCTGCACTTCAAGGAC	2198
DB	2095	CTGGAGAGAGCTGGGGGGACAGCTGATCCAGAGGCCACGGGTCTCTGCACTTCAAGGAC	2154
QY	2199	AGTTAACCAACACTGGCGCTATCCATCCAGATGTGCCAGCTCTCTGTGGAGAGATTAAG	2258
DB	2155	AGTTAACCAACACTGGCGCTATCCATCCAGATGTGCCAGCTCTCTGTGGAGAGATTAAG	2214
QY	2259	CTCCCTTGTGAGTACCGAGGAGTCCCTTTTATCACATCTGGAATGCACGCGAGCGGTAC	2318
DB	2215	CTCCCTTGTGAGTACCGAGGAGTCCCTTTTATCACATCTGGAATGCACGCGAGCGGTAC	2274
QY	2319	TTGCACTGCACCTTCCACCTGGAGGCTGTACGCCCCAGACCTAGTGACTTGGCTTGAAG	2378
DB	2275	TTGCACTGCACCTTCCACCTGGAGGCTGTACGCCCCAGACCTAGTGACTTGGCTTGAAG	2334
QY	2379	CTGTGGGTGTGCAGGTGGAGGGCGACGGGCGAGAGCTTTCAGCATCAACTTCAACATCACC	2438
DB	2335	CTGTGGGTGTGCAGGTGGAGGGCGACGGGCGAGAGCTTTCAGCATCAACTTCAACATCACC	2394

#### RESULT 4

КЕБОНІ 4  
UIS-10-311-623-13

US-10-311-623-13  
: Commence 13 Application US/10311623

Sequence 13, Application OS/103  
Publication No. US20040023244A1

; Publication No. US20  
GENERAL INFORMATION:

; GENERAL INFORMATION:  
ADDRESSEE: INCYTE GENOMICS INC : CPTFEIN Jennifer A.

APPLICANT: KATLYCK Deborah A : TRIBOULEY Catherine M.

APPLICANT: KALLICK, DEBORAH A.; TRIBOULET, DANIEL B.

; APPLICANT: YUE, Henry; NGUYEN, Daniel B.

; APPLICANT: TANG, Y. Tom; LAL, Preeti G.  
NOT TOWN: TANG, Y. Tom; LAL, Preeti G.  
NOT TOWN: TANG, Y. Tom; LAL, Preeti G.

APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda

APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.

APPLICANT: YAO, Monique G.; BUREFORD, Nell

; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mar

; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.

; APPLICANT: YANG, Junming; XU, Yuming

APPLICANT: GANDHI, Ameena R.; WARREN, Bridget

; APPLICANT: DING, Li; SANJANWALA, Madhu

APPLICANT: DUGGAN, Brendan M.

; TITLE OF INVENTION: RECEPTOR

; FILE REFERENCE: PF-0793 USN

CURRENT APPLICATION NUMBER: US/1

CURRENT FILING DATE: 2002-12-17

PRIOR APPLICATION NUMBER: US 0

PRIOR FILING DATE: 2001-06-21

PRIOR APPLICATION NUMBER: US 6,111,111

PRIOR FILING DATE: 2000-06-21

: PRIOR FILING DATE: 2000-03-21  
: PRIOR APPLICATION NUMBER: US 9

; PRIOR AFFILIATION NUMBER: 00 00 220 01  
 ; PRIOR FILING DATE: 2000-08-25

PRIOR FILING DATE: 2000-08-23  
PRIOR APPLICATION NUMBER: US

; PRIOR APPLICATION NUMBER  
: PRIOR FILING DATE: 2000-

; PRIOR FILING DATE: 2000  
 - NUMBER OF SEQ ID NOS: 3

; NUMBER OF S...

```

; SOFTWARE: PERI
CSC ID NO 13

```

SEQ ID NO	13
1	2500

; LENGTH: 3580

```

; TYPE: DNI

```

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: min

```

; OTHER INFORMATION: Incyte ID No. US20040023244A1 60523/1LCB1
US-10-311-623-13

Query Match      83.4%  Score 2402.6;  DB 17;  Length 3580;
Best Local Similarity 93.4%  Pred. No. 0;
Matches 2615;  Conservative 0;  Mismatches 9;  Indels 177;  Gaps 3;

QY      84  GCCATGCGCGTCGGCGCGCGCTGTGCCACGCGTCTCTGGGCATAGTCCTCGCGCGTTGG 143
          1  GCCATGCGCGTCGGCGCGCGCTGTGCCACGCGTCTCTGGGCATAGTCCTCGCGCGTTGG 60

```

QY 144 CTCGCGGCTCGGTGCGCCAGAGATGCCACGCTGGCCAAACCCAGTGCCTCGTGCAC 203  
DB |||||  
61 CTCGCGGCTCGGTGCGCCAGAGATGCCACGCTGGCCAAACCCAGTGCCTCGTGCAC 120  
QY 204 CCGGACCTGCTTCCCACTTCTCTGTGAGCCCGAGGATGTGTATCATGCTCAAGAACAG 263  
DB |||||  
121 CCGGACCTGCTTCCCACTTCTCTGTGAGCCCGAGGATGTGTATCATGCTCAAGAACAG 180  
QY 264 CCAGTGTCTGTGTGCAAGGCGCTGCGCCAGCAGATCTTCTTCAAGTGCACAGGG 323  
DB |||||  
181 CCAGTGTCTGTGTGCAAGGCGCTGCGCCAGCAGATCTTCTTCAAGTGCACAGGG 240  
QY 324 GAGTGGGTGCGCCAGGTGGACCAAGTATCGAGCGCAGCAGAGCGGAGTGGTGG 383  
DB |||||  
241 GAGTGGGTGCGCCAGGTGGACCAAGTATCGAGCGCAGCAGAGCGGAGTGGTGG 300  
QY 384 CCGACCATGAGAGTCCGCAATTAATGTCTCAAGGCGAGGTGCGAGAGTGTGGGCTG 443  
DB |||||  
301 CCCACCATGAGAGTCCGCAATTAATGTCTCAAGGCGAGGTGCGAGAGTGTGGGCTG 360  
QY 444 GAGGAATACTGCTGCGCATGCTGCGCATGAGTCTCTCGGCGACCCACCAAGTCAAG 503  
DB |||||  
361 GAGGAATACTGCTGCGCATGCTGCGCATGAGTCTCTCGGCGACCCACCAAGTCAAG 420  
QY 504 GCCTACATCCGATAGCGCATGCTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGG 563  
DB |||||  
421 GCCTACATCCGATAGCGCATGCTTTCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGG 480  
QY 564 GTGTCCCTGAGCAGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623  
DB |||||  
481 GTGTCCCTGAGCAGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 624 GAGGTGAGTGGTCCGGAACGAGACCTGTGTGACCCGCTGCTGCTGCTGCTGCTGCTG 683  
DB |||||  
541 GAGGTGAGTGGTCCGGAACGAGACCTGTGTGACCCGCTGCTGCTGCTGCTGCTGCTG 600  
QY 684 ATACCGGGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743  
DB |||||  
601 ATACCGGGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 744 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803  
DB |||||  
661 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
QY 804 TACGTGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 863  
DB |||||  
721 TAC-----723  
QY 864 CGCGGTGCGAGAAACGAGCGCGAGCTGCACCAACCCCGCGCTCTCTCAACGGGGGCGCT 923  
DB |||||  
724 -----723  
QY 924 TTCTGTGAGGGGAGAAATGTCCATGACCGACCGTCTCTCTGCTGCTGCTGCTGCTG 983  
DB |||||  
724 -----GTGGAC 729  
QY 984 GGCAGTGGAGCCCGTGGAGCAAGTGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1043  
DB |||||  
730 GGCAGTGGAGCCCGTGGAGCAAGTGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789  
QY 1044 AGCCGTGAGTGTCTGACCCAGCACCAGCGAGGGGAGGAGTGCAGGAGCACTGAC 1103  
DB |||||  
790 AGCCGTGAGTGTCTGACCCAGCACCAGCGAGGGGAGGAGTGCAGGAGCACTGAC 849  
QY 1104 CTGGACACCCGCACTGTATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163  
DB |||||  
850 CTGGACACCCGCACTGTATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909  
QY 1164 GTGGCCCTCTATGTGGGCTCATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1223  
DB |||||  
910 GTGGCCCTCTATGTGGGCTCATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969

QY 1224 CTCATCTCTGTTTATTCGCGGAAGAGAGGGGCTGGACTCAGATGTGTGCTGACTCGTCC 1283  
DB |||||  
970 CTCATCTCTGTTTATTCGCGGAAGAGAGGGGCTGGACTCAGATGTGTGCTGACTCGTCC 1029  
QY 1284 ATTCTCAGCTCAGGCTTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAGAG 1343  
DB |||||  
1030 ATTCTCAGCTCAGGCTTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAG 1089  
QY 1344 CTGCTCAGCTCAGGCTTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAG 1400  
DB |||||  
1090 CTGCTCAGCTCAGGCTTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAG 1149  
QY 1401 CCCCGGAGAGTGGGCTCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAGAG 1460  
DB |||||  
1150 CCCCGGAGAGTGGGCTCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAG 1209  
QY 1461 CTGCTGAGGCTTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAGAGAG 1520  
DB |||||  
1210 CTGCTGAGGCTTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAGAG 1269  
QY 1521 GTCTCCGCTCTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAGAGAG 1580  
DB |||||  
1270 GTCTCCGCTCTCCAGCCGCTCAGCTCAAGCCAGCAGCAAGAGAGAGAGAGAGAGAGAG 1329  
QY 1581 ACCTATGGGACCTTCAACTTCTCTCGGGGCGGCTGATGATGATGATGATGATGATGATGAT 1640  
DB |||||  
1330 ACCTATGGGACCTTCAACTTCTCTCGGGGCGGCTGATGATGATGATGATGATGATGATGAT 1389  
QY 1641 CTCCTCATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTATGAGATCTATGAGAT 1700  
DB |||||  
1390 CTCCTCATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGATCTATGAGATCTATGAGAT 1449  
QY 1701 CACAAGCCGGAAGAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1760  
DB |||||  
1450 CACAAGCCGGAAGAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1509  
QY 1761 GTTAGCTGTGGACCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1820  
DB |||||  
1510 GTTAGCTGTGGACCCCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1569  
QY 1821 TGTGGGAGGAGCCAGCTGACAGTGGAGCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1880  
DB |||||  
1570 TGTGGGAGGAGCCAGCTGACAGTGGAGCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1629  
QY 1881 AGCTGGAGAGAGATGTGTGCACTGGCGAGGAGCGGCTTGGCTTGGCTTGGCTTGGCTTGG 1940  
DB |||||  
1630 AGCTGGG---AGGATGTGTGACCTTGGCGAGGAGCGGCTTGGCTTGGCTTGGCTTGGCTTGG 1686  
QY 1941 CAGCTGGAGGAGGAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2000  
DB |||||  
1687 CAGCTGGAGGAGGAGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1746  
QY 2001 GGAGAGGCGCTCAGCGTGGCTTGGCGCAAGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2060  
DB |||||  
1747 GGAGAGGCGCTCAGCGTGGCTTGGCGCAAGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1806  
QY 2061 GCTGTGCACTTCCCTCGAGTACAACATCCGGGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 2120  
DB |||||  
1807 GCTGTGCACTTCCCTCGAGTACAACATCCGGGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 1866  
QY 2121 CTCAGGAGGAGTGGTGGAGAGAGAGTGGGGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2180  
DB |||||  
1867 CTCAGGAGGAGTGGTGGAGAGAGAGTGGGGGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1926  
QY 2181 GTCTGTGCACTTCAAGGAGAGTGGTGGAGAGAGTGGGGGAGAGTGGTGGTGGTGGTGGTGG 2240  
DB |||||  
1927 GTCTGTGCACTTCAAGGAGAGTGGTGGAGAGAGTGGGGGAGAGTGGTGGTGGTGGTGGTGG 1986  
QY 2241 TCCCTGTGGAGAGTAAAGTCTCTTGTGAGTACCAAGGAGATCCCTTTTATCACATCTGG 2300  
DB |||||  
1987 TCCCTGTGGAGAGTAAAGTCTCTTGTGAGTACCAAGGAGATCCCTTTTATCACATCTGG 2046  
QY 2301 AATGGCAGCAGCGGTAATTGCACTGCACTTCACTTGGAGCGGTGTCAGCGCCCGCACT 2360

2047 AATGGCAGCAGCGTACTTGGCTACGCTTACCTGAGGCTGTACGCCAGCACT 2106  
2361 AGTAGCTGGCTGCAAGCTGTGGGTGTGCGAGTGGAGGGCAGCGGAGAGCTTCAGC 2420  
2107 AGTGAGCTGGCTGCAAGCTGTGGGTGTGCGAGTGGAGGGCAGCGGAGAGCTTCAGC 2166  
2421 ATCAACTTCAACATCAACAGGACACAAAGTTTGTGAGTGTGCTGTGGAGAGTGAA 2480  
2167 ATCAACTTCAACATCAACAGGACACAAAGTTTGTGAGTGTGCTGTGGAGAGTGAA 2226  
2481 GCGGGGTCCAGCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCCTCATTCGSCAG 2540  
2227 GCGGGGTCCAGCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCCTCATTCGSCAG 2286  
2541 AAGATATTTCCAGCTGGACCCACACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2600  
2287 AAGATATTTCCAGCTGGACCCACACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCC 2346  
2601 CAGAAATCCACCTGGACAGCCATCTCAGCTTCTTGGCTTCCAGCCAGCCGACAGCC 2660  
2347 CAGAAATCCACCTGGACAGCCATCTCAGCTTCTTGGCTTCCAGCCAGCCGACAGCC 2406  
2661 ATGATCTCAACCTGTGGAGGCGGCACTTCCCAAGCGCAACCTCAGCCAGCTGGCT 2720  
2407 ATGATCTCAACCTGTGGAGGCGGCACTTCCCAAGCGCAACCTCAGCCAGCTGGCT 2466  
2721 GCACAGTGGCTGACTGGGCGGCGAGAGCTGGCTTTCACAGTGTGGAGGCTGAG 2780  
2467 GCACAGTGGCTGACTGGGCGGCGAGAGCTGGCTTTCACAGTGTGGAGGCTGAG 2526  
2781 TGCTGAGCGCGGCGAGGCGGCGAGAGCTGGCTTTCACAGTGTGGAGGCTGAG 2840  
2527 TGCTGAGCGCGGCGAGGCGGCGAGAGCTGGCTTTCACAGTGTGGAGGCTGAG 2586  
2841 AGGAGAGCGGCGAGGCGGCGGCTTCCCAAGCGGCGGAG 2881  
2587 AGGAGAGCGGCGAGGCGGCGGCTTCCCAAGCGGCGGAG 2627

RESULT 5

US-09-933-261-1  
; Sequence 1, Application US/09933261  
; Publication No. US2003040046A1  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsey  
; Masu, Masayuki  
; Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/933,261  
; FILING DATE: 20-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/808,982  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UC96-217  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3014 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-933-261-1

Query Match 77.6%; Score 2235.8; DB 10; Length 3014;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;  
QY 87 ATGCGCTGCGGCGCGGCTGTGGCGAGCGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 146  
Db 1 ATGCGCTGCGGCGCGGCTGTGGCGAGTGTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 60  
QY 147 CGCGGCTCGGGTCCCGAGCAGAGTGCCACCGTGGCCAAACCGAGTCTGTGCAACCCG 206  
Db 61 CGTGGTTTCGGTCCCGAGCAGAGTGCCACCGTGGCCAAACCGAGTCTGTGCAACCCG 120  
QY 207 GACCTGCTTCCCACTTCTCTGGTGGAGCGGAGTGTACATCGTCAAGAACCAAGCCA 266  
Db 121 GACCTGCTTCCCACTTCTCTGGTGGAGCGGAGTGTACATCGTCAAGAACCAAGCCA 180  
QY 267 GTGCTGCTTGTGTGCAAGGCGCTGCGCGCCACGAGATCTTCTCAAGTGCACCGGGAG 326  
Db 181 GTGTTGTGTGTGCAAGGCTGTGCTGCGCCACCGAGATCTTCTCAAGTGCACCGGGAA 240  
QY 327 TGGTGCGCCAGTGGACCACTGATCAGCGCAGACAGCGGAGGAGTGTGTGAGCGG 386  
Db 241 TGGTGCGCCAGTGGACCACTGATCAGCGCAGACCGAGCGGAGTGTGTGAGCGG 300  
QY 387 ACCATGAGGTCCGATTAATGTCTCAAGCAGCAGGTTCGAGAGAGTGTTCGGGCTGGAG 446  
Db 301 ACCATGAGGTCCGATTAATGTCTCAAGCAGCAGGTTCGAGAGAGTGTTCGGGCTGGAG 360  
QY 447 GAATACCTGTCAGTGTGGCTGAGTGTCTCTGGGCGACCCACCAAGAGTCAAGAGGCC 506  
Db 361 GAATACCTGTCAGTGTGGCTGAGTGTCTCTGGGCGACCCACCAAGAGTCAAGAGGCC 420  
QY 507 TACATCCGCTATGACGATTCGCAAGAACTTCAGCAGGAGGCGCTGGCCAGAGGAGTG 566  
Db 421 TACATCCGCTATGACGATTCGCAAGAACTTCAGCAGGAGGCGCTGGCCAGAGGAGTG 480  
QY 567 TCCCTGGAGCAGGCGATCGTCTGCGCTGCGTCCAGCGAGGCGATCCCTCCAGCGCGAG 626  
Db 481 TCCCTGGAGCAGGCGATCGTCTGCGCTGCGTCCAGCGAGGCGATCCCTCCAGCGCGAG 540  
QY 627 GTGAGTGGCTCCGGAACGAGGACCTGTGTGAGCCGCTCCCTGGACCCCAATGTATACATC 686  
Db 541 GTGAGTGGCTCCGGAACGAGGACCTGTGTGAGCCGCTCCCTGGACCCCAATGTATACATC 600  
QY 687 ACAGCGGAGCAGCGCTGTGTGAGCAGGCGCGCTGTGTGACACCGGCGCACTACACC 746  
Db 601 ACAGCGGAGCAGCGCTGTGTGAGCAGGCGCGCTGTGTGACACCGGCGCACTACACC 660  
QY 747 TCGCTGGCCAAAGAACATCGTGCAGCGTCCGCGAGCGCTCCGCTGTGTATCGCTCTAC 806  
Db 661 TGTGTGGCCAAAGAACATCGTGCAGCGCTCCGCGAGCAGCTCTGTGAGCGGTCAATGTTAT 720  
QY 807 GTGAACGGTGGGTGCTCGACGTGGACCGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGT 866  
Db 721 GTGAACGGTGGGTGCTCGACGTGGACCGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGT 780  
QY 867 GGCTGCGAGAAACGAGCGGAGTGTGACCAACCGGCGGCTCTCAACCGGCGGCTTTC 926



NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 BUSH STREET, SUITE 3200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/256,702  
 FILING DATE: 27-Sep-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/933,261  
 FILING DATE: 20-Aug-2001  
 APPLICATION NUMBER: 08/808,982  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OSMAN, RICHARD A.  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: UC96-217  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 343-4341  
 TELEFAX: (415) 343-4342  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3014 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-256-702-1

Query Match 77.6%; Score 2235.8; DB 15; Length 3014;  
 Best Local Similarity 88.5%; Pred. No. 0;  
 Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

Qy	87	ATGCGCGTCCGGCCCGGCGCTGTGGCAGTGTCTCTGGGCATAGTCTCTGGCGCTGGCTC	146
Db	1	ATGCGCGTCCGGCCCGGCGCTGTGGCAGTGTCTCTGGGCATAGTCTCTGGCGCTGGCTC	60
Qy	147	CGCGCGTCCGGTCCCGCAGCAGTGCACCGTGGCCAAACCCAGTGCCTGGTGCACAAACCG	206
Db	61	CGTGGTCCGGTCCCGCAGCAGTGCACCGTGGCCAAACCCAGTGCCTGGTGCACAAACCG	120
Qy	207	GACCTGTCTCCCACTTCTCTGGTGGAGCCCGAGGATGTATCGTCAAGAAACAGCCA	266
Db	121	GACCTGTCTCCCACTTCTCTGGTGGAGCCCGAGGATGTATCGTCAAGAAACAGCCA	180
Qy	267	GTGCTGTCTGGTGTGAAGCCGTCGCCCGCAGCAGATCTCTCAAGTGCAACCGGGAG	326
Db	181	GTGCTGTCTGGTGTGAAGCCGTCGCCCGCAGCAGATCTCTCAAGTGCAACCGGGAA	240
Qy	327	TGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACCGGAGCAGTGGTGGAGCCG	386
Db	241	TGGGTGGCCAGGTGGACCACTGATCGAGCGCAGCAGACCGGAGCAGTGGTGGAGCCG	300
Qy	387	ACCATGGAGTCCCGATTAATGTCTCAAGGAGCAGGTCGAGAGGTGTTCGGGCTGGAG	446
Db	301	ACCATGGAGTCCCGATTAATGTCTCAAGGAGCAGGTCGAGAGGTGTTCGGGCTGGAG	360
Qy	447	GAATAGTGTGCAGTGTGGCATGGAGTCTCTGGGACCAACCAAGATCAGAGGCC	506
Db	361	GAATAGTGTGCAGTGTGGCATGGAGTCTCTGGGACCAACCAAGATCAGAGGCC	420
Qy	507	TACATCCGATAGCCAGATTTGGCAGAGAACCTTCGAGCAGGACCCGCTGGCCAAAGAGGTG	566

Db	421	TACATCCGATTTGCTTAITTTGGCAAGAACTTTTGACAGAGCCACTGGCCAAAGAGTG	480
Qy	567	TCCCTGGACAGGGCATCTGTGTCCTGCCGTCACCGGAGGGCATCTCTCCAGCCGAG	626
Db	481	TCACTGGACAAAGGCATTGTACTCTTGTGCCCGCCAGAGGAATCCCCCAGTGTAG	540
Qy	627	GTGGAGTGGCTCCGGAACGAGGACCTGTGGAGCCCGTCCCTTGAGACCCCAATGTATCATC	686
Db	541	GTGGAGTGGCTTCGAATGAGGACCTGTGGAGCCCGTCCCTTCGATCCCAATGTATCATC	600
Qy	687	ACCGGGAGCAGACCTGTGTGTGACAGGCGCGCTTGTCTGACAGGCGCAACTACAC	746
Db	601	ACCGGGAGCAGACCTGTGTGTGACAGGCGCGCTTGTCTGACAGGCGCAACTACAC	660
Qy	747	TGCGTGGCCAAAGAACATCGTGGCACCTGCGCGCAGCGCTCCCGTGTGTCTCATGCTAC	806
Db	661	TGTGTGGCCAAAGAACATCGTAGCCCGTTCGCGGAAGACCTCTGAGGGGTCAITGTTAT	720
Qy	807	GTGAACGGTGGGTGGTGCAGCGTGCAGCGAGTGTGTGGTGTGACGCGCAGCTGTGGCGC	866
Db	721	GTGAACGGTGGGTGGTGCAGCGTGCAGCGAGTGTGTGGTGTGACGCGCAGCTGTGGCGC	780
Qy	867	GGCTGGCAGAAACGGAGCGCGAGCTGCACCAACCGCGGCTCTCAACGGGGCGCTTC	926
Db	781	GGCTGGCAGAAACGGAGCGCGAGCTGCACCAACCGCGGCTCTCAACGGGGCGCTTC	840
Qy	927	TGTGAGGGCAGAAATGTCATGACCGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	986
Db	841	TGTGAGGGCAGAAATGTCATGACCGCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	897
Qy	987	AGCTTGAGCCCGTGGAGCAAGTGTGTGGCTGTGGGTGTGACTGACACCACTGGCGGAGC	1046
Db	898	AGCTTGAGTTCGTGGAGTAAGTGTGTGGCTGTGGGTGTGACTGACACCACTGGCGGAGC	957
Qy	1047	CGTGTGTCTGTACCCAGCACCCCGCAACGGAGGGAGGAGTGCAGGCGCAGTGCCTG	1106
Db	958	CGCGAGTGTCTGTACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTG	1017
Qy	1107	GACACCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGTGGCCCTGAGGAGCTG	1166
Db	1018	GACACCGCAACTGTACAGTGAACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1077
Qy	1167	GCCCTCTATGTGGGCTCATCGCGTGTGGCGCTGTGGCTGTGGCTGTGGCTGTGGCTGT	1226
Db	1078	GCTCTCTACATCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1137
Qy	1227	ATCTCGTGTATTCGCGAAGAGAGGGGTGTGACTCAGATGTGGTGTGCTGTGCTCAAT	1286
Db	1138	GGACTCATTTACTGTGCAAGAGAGGGGTGTGACTCAGATGTGGCGGAGTGTGCTTCAAT	1197
Qy	1287	CTCACCTCAGGCTTCCAGCCCGTCAAGATCAAGCCCGAGCAAGAGAGAGCAACCCCACTG	1346
Db	1198	CTCACCTCAGGCTTCCAGCCCGTCAAGATCAAGCCCGAGCAAGAGAGAGCAACCCCACTG	1257
Qy	1347	CTCACCTCAGGCGGAGCTCAGACCACTCAGACCACTACCACTACCACTACCACTACCACT	1403
Db	1258	CTCACCTCAGGCGGAGCTCAGACCACTCAGACCACTACCACTACCACTACCACTACCACT	1317
Qy	1404	CGGAGGATGGCCCGGAGCCCAAGTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT	1463
Db	1318	AGGAGGATGGCCCGGAGCCCAAGTTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT	1377
Qy	1464	GGTGGCGCCGCCACACACTGACACACAGCTCTCCACCTCTGTGGCGGAGGAGTGTGTC	1523
Db	1378	GGGAGTGGCGGCATACCTTGACACACAGCTCACCCACTCTGTGGGCTGAGGACTTCGTC	1437
Qy	1524	TCCCGCTCTCCACCGAATCTATCTTCGCTTCCCTGCCCGGAGGAGCAGCAACATGAC	1583
Db	1438	TCCCGCTCTCCACCGAATCTATCTTCGCTTCCCTGCCCGGAGGAGCAGCAACATGAC	1497
Qy	1584	TATGGGACCTTCACTCTCTCGGGGGCGGCTGATGATCCCTTAATACAGGTATACGCTTC	1643
Db	1498	TACGGGACCTTCACTCTCTCGGGGGCGGCTGATGATCCCTTAATACGGGATACGCTTC	1557









QY 2784 TGA 2786  
 Db 2695 TGA 2697  
 RESULT 8  
 US-09-933-261-2  
 ; Sequence 2, Application US/099333261  
 ; Publication No. US2003004046A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tessier-Lavigne, Marc  
 ; Leonardo, E. David  
 ; Hink, Lindsay  
 ; Masu, Masayuki  
 ; Kazuko, Keino-Masu  
 ; TITLE OF INVENTION: Netrin Receptors  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 268 BUSH STREET, SUITE 3200  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94104  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/933,261  
 ; FILING DATE: 20-Aug-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/808,982  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OSMAN, RICHARD A  
 ; REGISTRATION NUMBER: 36,627  
 ; REFERENCE/DOCKET NUMBER: UC96-217  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 343-4341  
 ; TELEFAX: (415) 343-4342  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1787 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-933-261-2  
 Query Match 54.5%; Score 1570.4; DB 10; Length 1787;  
 Best Local Similarity 96.9%; Pred. No. 0;  
 Matches 1721; Conservative 0; Mismatches 36; Indels 19; Gaps 11;  
 QY 1114 GCAACTGTACCACTGCTGTGTACACAGTCTTCTGGCCCTGAGGACGTGGCCCTCT 1173  
 Db 1 GCAACTGTACCACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 59  
 QY 1174 ATGTGGGCTCATCGCGGTGGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1233  
 Db 60 ATGTGGGCTCATCGCGGTGGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 119  
 QY 1234 TTTATTGCCGGAAGAAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCAATTCTCACT 1293  
 Db 120 TTTATTGCCGGAAGAAGAGGGGCTGGACTCAGATGTGGCTGACTCGTCAATTCTCACT 179  
 QY 1294 CAGGCTTCCAGCCCTCAGCATC-AAGCCAGCAAGCAGAAAGCCCATCTGTCTACC 1352  
 Db 180 CAGGCTTCCAGCCCTCAGCATCTAAGCCAGCAAGCAGAAAGCCCATCTGTCTACC 239

QY 1353 ATCCAGCCGAGACCTCAG---CACCAACCACTACCAAGGCGAGTCTTGTCCCCGGCAG 1409  
 Db 240 ATCCAGCCGAGACCTCAGCAACCACTACCAAGGCGAGTCTTGTCCCCGGCAG 299  
 QY 1410 GATGGCCCGAGCCCAAGTTCCAGTTCACCAATGGGACCTGTCTAGCCCCCTGGGTGGC 1469  
 Db 300 GATGGCCCGAGCCCAAGTTCCAGTTCACCAATGGGACCTGTCTAGCCCCCTGGGTGGC 359  
 QY 1470 GCGCGCCACACACTGCACACAGCTCTCCCACTCTGAGGCGGAGAGTTGCTCTCCCGC 1529  
 Db 360 GCGCGCCACACACTGCACACAGCTCTCCCACTCTGAGGCGGAGAGTTGCTCTCCCGC 419  
 QY 1530 CTCTCCACCCAGAACTCTTCGGTTCCTGCCGAGGACCAAGCAATGACCTATGGG 1589  
 Db 420 CTCTCCACCCAGAACTCTTCGGTTCCTGCCGAGGACCAAGCAATGACCTATGGG 479  
 QY 1590 ACCTTCAACTTCTCTGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATC 1649  
 Db 480 ACCTTCAACTTCTCTGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATC 539  
 QY 1650 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTCACAAGCG 1709  
 Db 540 CCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTCACAAGCG 599  
 QY 1710 GAAGACGTGAGTTGCCCTAGTGGTGTGAGACCCCTGCTGAGTCCCATCTTAGTGT 1769  
 Db 600 GAAGACGTGAGTTGCCCTAGTGGTGTGAGACCCCTGCTGAGTCCCATCTTAGTGT 659  
 QY 1770 GGACCCCTGGCTGCTGCTCACCAGGCGAGTCTCTGGCTATGAGACCTGTGGGAG 1829  
 Db 660 GGACCCCTGGCTGCTGCTCACCAGGCGAGTCTCTGGCTATGAGACCTGTGGGAG 719  
 QY 1830 CCCAGCCCTGACAGCTGGAGCCTGGGCTCAAAAAGCAGTCGTCGAGGCGAGCTGGAG 1889  
 Db 720 CCCAGCCCTGACAGCTGGAGCCTGGGCTCAAAAAGCAGTCGTCGAGGAGG-AGCTGGAG 778  
 QY 1890 CAGGATGTGCTGACCTGGCGGAGGAGGCCCTCCACCTCTACTACTGCACTGGAG 1949  
 Db 779 GATG---TCTGCACCTGGCGGAGGAGGCCCTCCACCTCTACTACTGCACTGGAG 834  
 QY 1950 GCCAGTGCCTGCTACCTCTTACCGAGCAGCTGGCGGCTTTCCTGTGGGAGAGGCC 2009  
 Db 835 GCCAGTGCCTGCTACCTCTTACCGAGCAGCTGGCGGCTTTCCTGTGGGAGAGGCC 894  
 QY 2010 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTCTTCTGTTTGGCGGCTGCTGCACC 2069  
 Db 895 CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTCTTCTGTTTGGCGGCTGCTGCACC 954  
 QY 2070 TCCCTCGAGTACAAATCCGGGTCTACTGCTGATGACACCCACGATGCACCTCAAGGAG 2129  
 Db 955 TCCCTCGAGTACAAATCCGGGTCTACTGCTGATGACACCCACGATGCACCTCAAGGAG 1014  
 QY 2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCCAGGGTCTCTGCAC 2189  
 Db 1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCCAGGGTCTCTGCAC 1074  
 QY 2190 TTCAAGGACAGTTACCAACCTGGCGCTTATCCATCCAGATGTGCCAGCTCCCTGTGG 2249  
 Db 1075 TT-AAGGACAGTTACCAACCT--GCCCTATCATCCAGATGTGCCAGCTCCCTGTGG 1131  
 QY 2250 AAGAGTACCTCTTGTGCTAGCTTACAGGAGATCCCTTTTATCATCTGGAATGGCAG 2309  
 Db 1132 AAGAGTACCTCTTGTGCTAGCTTACAGGAGATCCCTTTTATCATCTGGAATGGCAG 1191  
 QY 2310 CAGCGGTACTTGCATGCACCTTCACTTGGAGGCTGTGAGCCCGCAGCTAGTACCTG 2369  
 Db 1192 CAGCGGTACTTGCATGCACCTTCACTTGGAGGCTGTGAGCCCGCAGCTAGTACCTG 1251  
 QY 2370 GCTTCAAGCTGTGGGTGTGGCAGGTGGGCGGAGGCGGAGCTTACGATCAACTTC 2429  
 Db 1252 GCTTCAAGCTGTGGGTGTGGCAGGTGGGCGGAGGCGGAGCTTACGATCAACTTC 1311



2010 CTACGCTGGTGGCGCAAGCGCTCAAGCTCTTCTGTTTGGCCGCTGCTGCACC 2069  
|  
895 CTCAGCTGGTGGCGCAAGCGCTCAAGCTCTTCTGTTTGGCCGCTGCTGCACC 954  
|  
2070 TCCCTCGAGTCAACATCCGGGTCTACTGCTGATGACACCCAGATGCACTCAAGAG 2129  
|  
955 TCCCTCGAGTCAACATCCGGGTCTACTGCTGATGACACCCAGATGCACTCAAGAG 1014  
|  
2130 GTGTGTCAGCTGGAGAACGCTGGGGAGACGATGATCCAGGAGCCAGGCTGTCGAC 2189  
|  
1015 GTGTGTCAGCTGGAGAACGCTGGGGAGACGATGATCCAGGAGCCAGGCTGTCGAC 1074  
|  
2190 TTCAGGACAGTTTACCAACCTTGGCTATCCATCCACGATGTCGCCAGCTCCCTGGG 2249  
|  
1075 TT-AAGGACAGTTTACCAACCT--GCCCTATCATCCAGATGTCGCCAGCTCCCTGGG 1131  
|  
2250 AAGAGTAACTCTTGTGACGTACAGGAGATCCCTTTTATCATCTGGAATGGCAGC 2309  
|  
1132 AAGAGTAACTCTTGTGACGTACAGGAGATCCCTTTTATCATCTGGAATGGCAGC 1191  
|  
2310 CAGCGGTACTTGACACTGCACCTTCCACCTGGAGGCTGTCAGCCCGAGCACTAGTACCTG 2369  
|  
1192 CAGCGGTACTTGACACTGCACCTTCCACCTGGAGGCTGTCAGCCCGAGCACTAGTACCTG 1251  
|  
2370 GCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTTCAGCATCAATTC 2429  
|  
1252 GCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTTCAGCATCAATTC 1311  
|  
2430 AACATCACCAGGACACAAAGTTTGTGAGTGTGCTGTGAGAGTGAAGCGGGGTC 2489  
|  
1312 AACATCACCAGGACACAAAGTTTGTGAGTGTGCTGTGAGAGTGAAGCGGGGTC 1371  
|  
2490 CCAGCCCTGTGGGCGCCAGTGTCTCAAGATCCCTTCTCATTCGGCAGAGATAATT 2549  
|  
1372 CCAGCCCTGTGGGCGCCAGTGTCTCAAGATCCCTTCTCATTCGGCAGAGATAATT 1431  
|  
2550 TCCAGCTGTGACCCACCTTGTAGGCGGGGTGCGGACTTGGCGGACTCTGGGCCAGAACTC 2609  
|  
1432 TCCAGCTGTGACCCACCTTGTAGGCGGGGTGCGGACTTGGCGGACTCTGGGCCAGAACTC 1491  
|  
2610 CACTTGACAGCCATCTCAGCTTCTTTGCTCTCAAGCCAGCCCGACAGCATGATCCTC 2669  
|  
1492 CACTTGACAGCCATCTCAGCTTCTTTGCTCTCAAGCCAGCCCGACAGCATGATCCTC 1551  
|  
2670 AACCTGTGGAGCGCGGCACTTCCCAACGGCACTCAGCCAGCTGGCTGACAGCTG 2729  
|  
1552 AACCTGTGGAGCGCGGCACTTCCCAACGGCACTCAGCCAGCTGGCTGACAGCTG 1611  
|  
2730 GTGGACTGGGCGGCGAGACGCTGGCCCTC-TTCACAGTG-TGGAGGCTGAGTGTGAG 2787  
|  
1612 GTGGAGCTGGGCGGCGAGACGCTGGCCCTC-TTCACAGTG-TGGAGGCTGAGTGTGAG 1671  
|  
2788 GCCGCGAGCGCGGACACCTACACTCTACAGCTTTGGCA--CCACCAAGGACAGGCA 2845  
|  
1672 GCCGCGCAGG--CGAACACTTCAAAATTTTACAGTTTGGGAAACCCACCAAGGACAGGCA 1729  
|  
2846 GAAGCGGACAGCGGCGCTTCCCAACACCGGGGAGA 2881  
|  
1730 GAAGCGGACAGCGGCTTTTCCAAACCGGGGAGA 1765  
|

## RESULT 10

US-10-296-115-365  
; Sequence 365, Application US/10296115  
; Publication No. US20040053248A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq Inc  
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 784PCT  
; CURRENT APPLICATION NUMBER: US/10/296,115  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 1478

; SEQ ID NO 365  
; LENGTH: 1321  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-296-115-365

Query Match 41.4%; Score 1193.6; DB 13; Length 1321;  
Best Local Similarity 97.8%; Pred. No. 2.6e-287;  
Matches 1295; Conservative 0; Mismatches 19; Indels 10; Gaps 8;  
QY 1476 CACACACTGACACACAGCTCTCCACCTCTGAGCGCGAGGAGTTCTCCCGCTCTCC 1535  
|  
Db 1 CACACACTGACACACAGCTCTCCACCTCTGAGCGCGAGGAGTTCTCTCCCGCTCTCC 60  
|  
QY 1536 ACCAGAACTACTTCCGCTCTCCCGCGAGGACACAGCAACATGACTATGGACCTTC 1595  
|  
Db 61 ACCAGAACTACTTCCGCTCTCCCGCGAGGACACAGCAACATGACTATGGACCTTC 120  
|  
QY 1596 AACTTCTCCGGGCGCGCTGATGATCCCTAAATACAGGTATCAGCTCTCTATCCCCCA 1655  
|  
Db 121 AACTTCTCCGGGCGCGCTGATGATCCCTAAATACAGGTATCAGCTCTCTATCCCCCA 180  
|  
QY 1656 GATGCCATACCCCGAGGAAAGATCTATGAGATCTACTCAGCTGCAACAAGCGGAGAC 1715  
|  
Db 181 GATGCCATACCCCGAGGAAAGATCTATGAGATCTACTCAGCTGCAACAAGCGGAGAC 240  
|  
QY 1716 GTGAGTTGCCCTAGCTGTGCTGACACCTGCTGAGTCCCATGCTTACTGTGGACCC 1775  
|  
Db 241 GTGAGTTGCCCTAGCTGTGCTGACACCTGCTGAGTCCCATGCTTACTGTGGACCC 300  
|  
QY 1776 CCT-GGCGTCTCTCTCACCAGCGGCGAGTCACTCT-GGCTATGGACCACTGT-GGGAGCCC 1832  
|  
Db 301 CCTGGGCTCTCTCTTACCGCGCGAGTCACTCTGGGGATGAGACCACTGTGGGAGCCC 360  
|  
QY 1833 AGCCCTGACAGCT-GGAGCCTGCGCTCAAAAAGCAGTCTGCGAGGCGAGCTGGGAGCA 1891  
|  
Db 361 AGCCCTGACAGCTGGGAGCCTGCGCTCAAAAAGCAGTCTGCGAGGCGAGCTGGG--A 417  
|  
QY 1892 GATGTGCTGCACTGGGCGAGAGCGGCTCTCCACCTCTACTACTGCGAGCTGGAGGC 1951  
|  
Db 418 GATGTGCTGCACTGGGCGAGAGCGGCTCTCCACCTCTACTACTGCGAGCTGGAGGC 477  
|  
QY 1952 CAGTGTCTGTAGCTTCCACGAGCAGCTGGGCGGCTTTGGCCCTGGTGGGAGAGCCCT 2011  
|  
Db 478 CAGTGTCTGTAGCTTCCACGAGCAGCTGAGCGGCTATGGCCCTGGTGGGAGAGCCCT 537  
|  
QY 2012 CAGCGTGGTGGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGGCTGGCTGCACTTC 2071  
|  
Db 538 CAGCGTGGTGGCGCAAGCGCTCAAGCTGCTTCTGTTTGGCGGCTGGCTGCACTTC 597  
|  
QY 2072 CTTGAGTAAACATCCGGGTCTACTGCTGATGACACCCAGATGCACTCAAGAGGT 2131  
|  
Db 598 CTTGAGTAAACATCTACTGCTGATGACACTCAGATGCACTCAAGTCACTCAAGTACT 657  
|  
QY 2132 GGTGAGCTGGAGAGCAGCTGGGCGGACAGTATCCATCCAGATGTGCCAGCTCCCTTGGAA 2251  
|  
Db 658 GGTGAGCTGGAGAGCAGCTGAGGAGCAGCTGATCCAGGAGCCACGGTCTCGACTT 2191  
|  
QY 2192 CAAAGACAGTTACCAACCTTCCGCTTATCCATCCAGATGTGCCAGCTCCCTGTGAA 2251  
|  
Db 718 CAAAGACAGTTACCAACCTTCCGCTTATCCATCCAGATGTGCCAGCTCCCTGTGAA 777  
|  
QY 2252 GAGTAACTCTTGTGACGTACAGGAGATCCCTTTTATCACTCTGGAATGGACGCA 2311  
|  
Db 778 GAGTAACTCTTGTGACGTACAGGAGATCCCTTTTATCACTCTGGAATGGACGCA 837  
|  
QY 2312 CGGTTACTTGCAGTGCACCTTCCAGCTGGAGGCTGAGCCCGAGCACTAGTACCTGGC 2371  
|  
Db 838 CGGTTACTTGCAGTGCACCTTCCAGCTGGAGGCTGAGCCCGAGCACTAGTACCTGGC 897  
|

```

QY 2372 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCGAGAGCTTCAGCATCAACTTCAA 2431
Db 898 CTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCGAGAGCTTCAGCATCAACTTCAA 957

QY 2432 CATCACCAAGGACACAAAGTTTCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTGCC 2491
Db 958 CATCACCAAGGACACAAAGTTTCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTGCC 1017

QY 2492 AGCCCTGTGGGCCCGCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGCAGAGAATAATTTC 2551
Db 1018 AGCCCTGTGGGCCCGCCAGTGCCTTCAAGATCCCTTCTCTCATTCGGCAGAGAATAATTTC 1077

QY 2552 CAGCCTGGACCCACCTGTAGGGGGGTGCCGACTGGCGACTCTGGCCCGAGAACTCCA 2611
Db 1078 CAGCCTGGACCCACCTGTAGGGGGGTGCCGACTGGCGACTCTGGCCCGAGAACTCCA 1137

QY 2612 CTGGGACAGCATCTCAGCTTCTTTGCTCTCAAGCCCGAGCCCGCCAGCCATGATCCTCAA 2671
Db 1138 CTGGGACAGCATCTCAGCTTCTTTGCTCTCAAGCCCGAGCCCGCCAGCCATGATCCTCAA 1197

QY 2672 CTTGTGGAGGCGGGCAGCTTCCCAACGGCAACTCAGCCAGCTGGCTCAGCAGTGGC 2731
Db 1198 CTTGTGGAGGCGGGCAGCTTCCCAACGGCAACTCAGCCAGCTGGCTCAGCAGTGGC 1257

QY 2732 T-GGACTGGCCAGCCAGAGAGCTGGCCTC-TTCACAGTG-TCCGAGGCTGAGTCTGAGG 2788
Db 1258 TGGAGCTGGCCAGCAGGACGGTGGCTTCTTTACAGTGTTCGGAGGCTGAGTCTGAGG 1317

QY 2789 CCGG 2792
Db 1318 CCGG 1321

RESULT 11
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

Query Match 39.8%; Score 1147.8; DB 9; Length 1282;
Best Local Similarity 97.7%; Pred. No. 6.7e-276;
Matches 1259; Conservative 0; Mismatches 18; Indels 11; Gaps 9;

QY 1510 CCGAGAGTGTGTCGCCGCTCTCCACCCAGAACTACTTCGGCTCCCTGCCCGCCGAGCA 1569
Db 1 CCGAGAGTGTGTCGCCGCTCTCCACCCAGAACTACTTCGGCTCCCTGCCCGCCGAGCA 60

QY 1570 CCAGCAATGACCTATGGACCTTCACTTCTCGGGGGCGGCTGATGATCCCTTAATA 1629
Db 61 CCAGCAATGACCTATGGACCTTCACTTCTCGGGGGCGGCTGATGATCCCTTAATA 120

QY 1630 CAGGTATCAGCTCTCTATCCCGCCAGATGCCATACCCCGAGGGAGATCTATGATCT 1689
Db 121 CAGGTATCAGCTCTCTATCCCGCCAGATGCCATACCCCGAGGGAGATCTATGATCT 180

```

```

QY 1690 AACTCAGCTGCACAGCCGGAAGAGCTGAGGTTGCCCTAGCTGGCTGTGAGACCTGTC 1749
Db 181 AACTCAGCTGCACAGCCGGAAGAGCTGAGGTTGCCCTAGCTGGCTGTGAGACCTGTC 240

QY 1750 TGAATCCCATCGTTAGCTGTGACACCCCT - GGCGTCTCTGCTACCCGGGCGAGTATCCT - 1807
Db 241 TGAATCCCATCGTTAGCTGTGACACCCCTGGGGCTCTGCTACCCGGGCGAGTATCCTG 300

QY 1808 GGCTATGACCACTGTGGGAGCCCAAGCCCTGACAGCT - GGAGCTGCGCTCAAAAGC 1866
Db 301 GGCTATGACCACTGTGGGAGCCCAAGCCCTGACAGCTGGAGCTGCGCTCAAAAGC 360

QY 1867 AGTCGTGGGAGGCGAGCTGGGAGCAGATGCTGACCTGGGCGAGAGCGCCCTCC 1926
Db 361 AGTCGTGGGAGGCGAGCTGGGAGGATGTGCT - TGACCTGGGCGAGAGCGCCCTCC 418

QY 1927 ACCTTACTACTGCCAGCTGGAGGCCAGTCTGCTACGTCTTCAACGAGAGTGGGCC 1986
Db 419 ACCTTACTACTGCCAGCTGGAGGCCAGTCTGCTACGTCTTCAACGAGAGTGGGCC 478

QY 1987 GCTTGGCCCTGGTGGGAGGCGCTCAGCGTGGCTGCCCAAGCGCTCAAGCTGCTTC 2046
Db 479 GCTTGGCCCTGGTGGGAGGCGCTCAGCGTGGCTGCCCAAGCGCTCAAGCTGCTTC 538

QY 2047 TGTGCGCGGTGGCTGACCTCCCTCGAGTACAACATCCCGGTCTACTGCTGCTGATG 2106
Db 539 TGTGCGCGGTGGCTGACCTCCCTCGAGTACAACATCTGCTGCTGCTGCTGATG 598

QY 2107 ACACCCACGATGACCTCAAGGAGGTGCTGAGCTGGAGAGAGCTGGGGGAGCAGCTGA 2166
Db 599 ACACCTACGATGACCTCAAGTGTGCTGAGAGAGCTGGAGAGAGCTGGGGGAGCAGCTGA 658

QY 2167 TCAGGAGCCACGGCTCTGCACTTCAAGGACAGTGTACCAACCTCGCCCTATCCATCC 2226
Db 659 TCAGGAGCCACGGCTCTGCACTTCAAGGACAGTGTACCAACCTCGCCCTATCCATCC 718

QY 2227 ACGATGTCGCAGCTCCCTGTGGAAGAGTAAAGTCTTGTGCTGAGTACACAGAGATCCCT 2286
Db 719 ACGATGTCGCAGCTCCCTGTGGAAGAGTAAAGTCTTGTGCTGAGTACACAGAGATCCCT 778

QY 2287 TTTATCATCTGGAATGGCACGCGGTACTTGCATGCACTTCAACCTTCAACCTGAGCGGTG 2346
Db 779 TTTATCATCTGGAATGGCACGCGGTACTTGCATGCACTTCAACCTTCAACCTGAGCGGTG 838

QY 2347 TCAGCCCGCAGCTAGTGAACCTGCGCTGCAAGCTGTGGGTGGCAGGTGAGGGCGAGC 2406
Db 839 TCAGCCCGCAGCTAGTGAACCTGCGCTGCAAGCTGTGGGTGGCAGGTGAGGGCGAGC 898

QY 2407 GGCAGAGCTTCAGCATCAACTTCAACATCAACAGGACACAAGGTTTGTGAGTGTGCG 2466
Db 899 GGCAGAGCTTCAGCATCAACTTCAACATCAACAGGACACAAGGTTTGTGAGTGTGCG 958

QY 2467 CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCT 2526
Db 959 CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCT 1018

QY 2527 TCCTCATTCGCGAAGATGA - TTTCCAGCTGGACCCACCTGTAGCGGGGTGCCGAC 2585
Db 1019 TCCTCATTCGCGAAGATGAATTTTCCAGCTGGACCCACCTGTAGCGGGGTGCCGAC 1078

QY 2586 TGGCGG - ACTCTGGCCCCAGAAATCCACTCGACAGCCATCTCAGCTTCTTTGCTCCAA 2644
Db 1079 TGGCGGAATCTGGCCCCAGAAATCCACTCGACAGCCATCTCAGCTTCTTTGCTCCAA 1138

QY 2645 GCCAGCCCCCAGCCATGATCTCAACTGTGGAGGCGGCACTTCCCAAGGGCAA 2704
Db 1139 GCCAGCCCCCAGCCATGATCTCAACTGTGGAGGCGGCACTTCCCAAGGGCAA 1198

QY 2705 CTTAGCCAGCTGGCTGAGCAGTGGCTGAGTGGGCGCAGCAGCTGGCTTCTTAC 2764
Db 1199 CTTAGCCAGCTGG - TGCAGCAGTGGCTGAGTGGG - AGCAGAGCTGG - CTCTTAC 1254

```



Db 1365 GGCCTGGGATACCGGAGACCGTGTATGCTGAGGACTCCACGACAAATCCCCA 1424  
 QY 1435 TCACCAATGGGACCTGCTCAGCCCC-----1460  
 Db 1425 TGACCAACTCTCTCTGCTGGACCCCTTACCAGCCTTAAGGTCAAGGTCTACAGCTCA 1484  
 QY 1461 -----CTGGGTGGGGCGGCC 1476  
 Db 1485 GCACACGGGCTCTGGGCCAGGCTGGCAGATGGGCTGACCTGCTGGGGGTCTTCCGCG 1544  
 QY 1477 ACACACTGCACACAGCTCTCCACCTCTGAGCGGAGGATGCTCTCCCGCTCTCCA 1536  
 Db 1545 CTGGACATACCTTAGGATTTTCGCCGGGACACCCACTTCTCTGACCTGCGGAGGCCA 1604  
 QY 1537 CCAGAACTAC-----TTCGGCTCCCTGCGCGAGSCACAGCAACATGACCTATGGGA 1590  
 Db 1605 GCTCGGTTCCAGCAGCTCTTGGGCTGCCCCGAGACCCAGGAGCAGCGTCAGGGGCA 1664  
 QY 1591 CTTCAACTTCTCGGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCTCATCC 1650  
 Db 1665 CTTTGGCTGCTCGGGTGGGAGGCTCAGCATCCCGGCAAGGGGTGAGCTTGTGGTGC 1724  
 QY 1651 CCCAGATGCATACCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCGG 1710  
 Db 1725 CCAATGAGCCATTTCCCGAGGCAAGTTCTACGAGATGATCTACTCATCAAGGCGAG 1784  
 QY 1711 AAGACGTGAGGTTGCCCTAGCTGGCTGTGAGCCCTGCTGAGTCCCATCGTTAGCTGTG 1770  
 Db 1785 AAGTACCTTCCGCTTTTCAAGAGGACCCAGACAGTATTGAGCCCTCGGTGACCTGTG 1844  
 QY 1771 GACCCCTGGGCTGCTGCTACCCGGCAGTATCTCTGGCTTGGATGACACTGTGGGAGC 1830  
 Db 1845 GACCCACAGGCTCTGCTGTGCGGCCCGCTCATCTTCAATGCCCTGCTGCTGCCAAG 1904  
 QY 1831 CCAGCCCTGACAGCTGGAGCTGCGCTCAAAAGCAGTCTGTGGAGGCGAGCTGGGAGC 1890  
 Db 1905 TCAGTGGCGGTGATGATCTTTTCAAGACCCAGGAGCCACAGGCGACCTGGGAG- 1963  
 QY 1891 AGGATGTGCTGACCTGGGCGAGGAGGCGCCCTCCACCTCTACTATGCGAGCTGGAGG 1950  
 Db 1964 --GAGGTGTTGACCTGGATGAGGAGACCCCTGAACACACCTGCTACTGCGAGCTGGAGC 2021  
 QY 1951 CCAGTGGCTGCTGCTCTTACCGAGAGCTGGGCGGCTTTGCGCTGTTGGAGGAGGCC 2010  
 Db 2022 CCAGGGCTGTGCATCTCTGTGACACAGCTGGGACCCCTAGTGTTCAGGGCGAGTCT 2081  
 QY 2011 TCAGCTGGCTGCGCCCAAGCGCTCTCAAGCTGCTTCTGTTGCGCGGCTGGCTGCACCT 2070  
 Db 2082 ATTCCGCTCAGCAGTCAAGCGGCTCAGCTGGCGGCTTTCGCCCCCGCTCTGCACCT 2141  
 QY 2071 CCCTCGAGTACAACATCCGGGTCTACTGCTGCTGATGACACCCAGATGATCAAGGAGG 2130  
 Db 2142 CCCTGGAGTACAGCTCTCCGGGTCTACTGCTGGAGACACGCTGTAGCAGTGAAGGAG 2201  
 QY 2131 TGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGAGCCACGGTCTGCACT 2190  
 Db 2202 TGCTGGAGCTGGAGCGGAGCTCTGGGGCGGATCTTGGTGGAGGAGCGGAAACCGCTAATGT 2261  
 QY 2191 TCAAGCAGATTACACACACCTGCGCTTATCCATCCAGGATGTGCCAGCTCCCTGTGGA 2250  
 Db 2262 TCAAGCAGATTACACACACCTGCGCTTATCCCTCATGACCTCCCGCCCATTTGGA 2321  
 QY 2251 AGAGTAAGCTCTGCTGAGTACACAGGAGATCCCTTTTATCAGATCTGGAATGCGAGC 2310  
 Db 2322 GGAGCAAGCTGTGGCCAAATACAGAGAGATCCCTTCTATCAGATTTGGAGTGGCAGCC 2381  
 QY 2311 ACCGTACTTGCAGCTGACCTTCACTGAGGCTGTCAGCCCGGAGTGTGAGCTAGTGCCTGG 2370  
 Db 2382 AGAAGGCCCTCCAGTGCATTTTACCCCTGGAGGACAGCTTGGCTTCCACAGAGCTCA 2441  
 QY 2371 CTTGCAAGCTGTGGGTGTGGCGAGGTGGAGGCGAGCGGCGAGAGCTTTCAGCATCACTTCA 2430

Db 2442 CCTGCAAGATCTCGTGGGCAAGTGGAGGGAGGGCCAGATATTCTCAGCTGCATACCA 2501  
 QY 2431 ACATCACCAAG--GACACAAGTTTGTGAGTGTGCTGTGCTCTGGAGAGTGAAGCGGGG 2487  
 Db 2502 CTTCTGGCAGAGACACTGTGCTCCCTGGACACTCTCTGCTCTGCCCCCTGGCAGCACTG 2561  
 QY 2488 TCCAGCCCTGTGGGCCCCAGTGGCTTCAAGATGCCCTTCTCTCATTTGGCAGAGATAA 2547  
 Db 2562 TCACCAACCAGCTGGGACCTTATGCTTCAAGATGCCACTGTCTCATCCGCGCAGAATAT 2621  
 QY 2548 TTTTCAGCTGTGACCCACCTCTAGCGGGGTGCCGACTGGCGGACTCTGGCCCCAGAAAC 2607  
 Db 2622 GCAACAGCTTAGTAGTCCCCCACTCACGGGGCAATGACTGGCGGATGTTAGCACAAGC 2681  
 QY 2608 TCCACCTGGACAGCACTCTCAGCTTCTTGGCTTCCAAAGCCAGCCGCCACAGCCATGATCC 2667  
 Db 2682 TCTCTATGACCGGTACTTGAATTACTTTGGCACCAAGAGCGAGCCCGGCTGTATCC 2741  
 QY 2668 TCAACCTGTGGGAGGCGGCACTTCCCCAAACGGCAACCTCAGCCAGCTGGCTGCAGCAG 2727  
 Db 2742 TGGACCTCTGGGAAGCTCTGACAGGAGCCATGGGAGCTTCAACAGCTGGCGAGTSCCT 2801  
 QY 2728 TGGCTGAGCTGGGCGAGCCAGAGCTGGCTCTTTCAGGTGTCGAGGCTGAGTGTCTGA 2786  
 Db 2802 TGGAGGAGATGGGCAAGAGTGAATGCTGTGTGCTGTGGCCACCGAGCGGGAGTGTCTGA 2860

RESULT 13

US-10-218-779-1  
 ; Sequence 1, Application US/10218779  
 ; Publication No. US2004002922A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Edinger, Shlomit  
 ; APPLICANT: MacDougall, John  
 ; APPLICANT: Millet, Isabelle  
 ; APPLICANT: Ellerman, Karen  
 ; APPLICANT: Stone, David  
 ; APPLICANT: Gerlach, Valerie  
 ; APPLICANT: Grosse, William  
 ; APPLICANT: Alsbrook II, John  
 ; APPLICANT: Lepley, Denise  
 ; APPLICANT: Rieger, Daniel  
 ; APPLICANT: Burgess, Catherine  
 ; APPLICANT: Casman, Stacie  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Boldog, Ferenc  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Padigaru, Muralidhara  
 ; APPLICANT: Mishra, Vishnu  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Shenoy, Suresh  
 ; APPLICANT: Rastelli, Luca  
 ; APPLICANT: Tchernev, Velizar  
 ; APPLICANT: Vernet, Corine  
 ; APPLICANT: Zerhusen, Bryan  
 ; APPLICANT: Malyankar, Uriel  
 ; APPLICANT: Guo, Xiaojia  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Gangolli, Esha  
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-214  
 ; CURRENT APPLICATION NUMBER: US/10/218,779  
 ; CURRENT FILING DATE: 2002-08-14  
 ; PRIOR APPLICATION NUMBER: 60/253,834  
 ; PRIOR FILING DATE: 2000-11-29  
 ; PRIOR APPLICATION NUMBER: 60/250,-926  
 ; PRIOR FILING DATE: 2000-11-30  
 ; PRIOR APPLICATION NUMBER: 60/264,180  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 60/313,656  
 ; PRIOR FILING DATE: 2001-08-20  
 ; PRIOR APPLICATION NUMBER: 60/327,456  
 ; PRIOR FILING DATE: 2001-10-05

```
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-218-779-1

Query Match      30.9%; Score 889; DB 13; Length 2860;
Best Local Similarity 60.7%; Pred. No. 2,1e-211;
Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;

QY 184 ACCAGTGCCTGGTGCCCAACCGGACCTGCTTCCCACTTCTGTGTGAGCCCGAGGATG 243
Db 168 ACTCCTTCCCGTCCGCGCCAGCAGAGCCGCTGCCCTACTTCTGCGAGGCCACAGGAGC 227
QY 244 TGTATATCGTCAAGAACAGCAGCTGCTGTGTGTCAGAGCCGTCGCCGCGACGACGA 303
Db 228 CTTACATTGTGAAGAACAGCCTGTGGAGCTCCGCTGCGCGGCCCTTCCCGCCACACAGA 287
QY 304 TCTTCTTCAAGTGCAACCGGGAGTGGGTGCGCCAGGTGGACACCGTGTATCGAGCGCAGCA 363
Db 288 TCTACTTCAAGTGAACCGGCGAGTGGGTACGCGAAGACGACACGTCACACAGGAAGGCC 347
QY 364 CAGACGGGAGCAGTGGTGAGCCGACCAATGGAGTCCGCAATTAATGTCTCAAGSGCAGCAGG 423
Db 348 TGGATGAGGCCACCGGCTCGCGGTGCGCGAGGTGCAGATCGAGGTGTGCGCGCAGCAGG 407
QY 424 TCGAAGAGGTGTTTCGGGCTGGAGGAATACTGTTCCAGTGGTGGCATGGAGCTTCCTCGG 483
Db 408 TGGAGGAGCTCTTTGGGCTGGAGGATTAATGGTCCAGTGGCTGGAGTCCCGCGG 467
QY 484 GCACCAACAGAGTCAAGAGCCCTACATCCGCATAGCCAGATTGCGCAAGAACTTTCGAGC 543
Db 468 GCACCAACAGAGTCCGCGAGCTACTGTCGCATCGCTACTCTCGGCAAGAACTTTCGATC 527
QY 544 AGAGCCGCTGGCCAAAGAGGTGTCCTGGAGCAGGCGATGCTGCTCCCTCGCGTCCAC 603
Db 528 AGGAGCCTCTGGGCAAGAGGTGCCCTGGACCATGAGGTTCCTTCAGTGGCGCCCGC 587
QY 604 CGAGGSGCATCCCTCCAGCGAGGTGAGTGGCTCCGGAACGAGGACCTGGTGAGACCCGT 663
Db 588 CGAGGSGGTGCTGTGGCGGAGGTGGAATGGGTCAAGAAATGAGGATGTATCGACGCCCA 647
QY 664 CCTTGAACCCCAATGTATATACATCAGCGGAGCAGACGCTGTGGTGGCAGCGCCCGC 723
Db 648 CCCAGGACACCAACTTCTCTCCTCACCATCGACACACCTCATCATCCGCGAGCCGCGC 707
QY 724 TTGCTGACACGGCCAACTACCTGCTGGCCCAAGAACATCGTGGCAGCTGCGCGCAGCG 783
Db 708 TGTGAGACACTGCAACTATACCTGCTGGCCCAAGAACATCGTGGCCAAACGCGCGAGCA 767
QY 784 CTTCCGCTGTGTCATGCTCTAGTCAACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 843
Db 768 CCACTGCCACCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 827
QY 844 TCTGACGCGCCAGCTGTGGCGCGGCTGGCAGAAACGAGCGCGGAGCTGCAACCAACCCCG 903
Db 828 CTTGCTCCAAACCGCTGTGGCGGAGGTGGCAGAGCGCACCCGAGCCTGCAACCAACCCCG 887
QY 904 CGGCTCTCAACGGGGCGCTTTCTGTGAGGGGCGAGAAATGTCCATGACCGCAGCCGCTCTCT 963
Db 888 CTTCACTCAACGAGGGGCGCTTCTGCGAGGGCCAGGCAATTCAGAA--GACCGCTGCA 944
QY 964 CTTGCTGTCTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGGTGGGCTGTGGGCTGTGGG 1023
Db 945 CCACCAATCTGCCAGTTCAGTGGGGCGTGGACGAGTGGAGCAAGTGGTCAAGCTGCGAGCA 1004
QY 1024 TGGACTGCAACCCACTGGCGGAGCCGTGAGTGTCTTGACCCAGCACCCCGCAACGAGGGG 1083
Db 1005 CTGAGTGTGCCCACTGGCGGTAGCCCGGAGTGTGATGGCGGCCCCACCCCAAGACGAGGCC 1064
QY 1084 AGGAGTGCCAGGGCACTGACTGGACACCCGCAACTGTGTACAGTGCACCTCTGTGTACACA 1143
```

```
Db 1065 GTGACTGCGAGCGGACGCTGTGTGGACTCTAAGAACTGCACAGATGGGTGTGCATGCAAC 1124
QY 1144 GTGCTTCTGGCCCTGAGGACGTGGGCCCTCTATGTGGGCTCATCGCCGTGGCCGCTGCC 1203
Db 1125 TGGAGGCTCAGGGGATGCGGCGTGTATGCGGGCTGTGGTGGCCATCTTCTGTGGTGG 1184
QY 1204 TGTCTCTGCTGCTTGTCTCTCATCTCTGTTATTTGCCGGAAGAGGAGGCTGGACT 1263
Db 1185 TGGCAATCTCATGCGGTGGGGTGGTGTACCGCGCAACTGCGCTGACTTCGACTCGACA 1244
QY 1264 CAGATGTGCTGACTTCCTCCATT---CTCACCTCAGGCTTTCAGCCCGTTCAGACTCAAG 1320
Db 1245 CAGACATCAGTACTCATCTCTGCTGCCCTGACTGTGGTGTTCACCCCGTCAACTTAAAGA 1304
QY 1321 CCAGAAAGCAGACAAACCCCATCTGCT-----CACCATCCAGCGGACCTCAGCACCA 1374
Db 1305 CGGCAAGGCCAGTAAACCGCAGCTCTACACCCCTCTGTGCTCTGACCTGACGAGCA 1364
QY 1375 CCACCACTTACAGGGAGTCTCTGTCCCGGAGGATGGGCCCGAGCCCAAGTTCCAGC 1434
Db 1365 GCGCGGCACTTACCGGAGACCCGCTGTATGCCCTGCAGGACTCCACCGACAAATCCCCA 1424
QY 1435 TCACCAATGGCACCTCTCTCAGCCCC-----TGAGCTTAAAGTCAAGTCTTACAGCTCCA 1460
Db 1425 TGACCAACTCTCTCTGTGACCCCTTACCCAGCTTAAAGTCAAGTCTTACAGCTCCA 1484
QY 1461 -----CTGGGTGGCGGCGCC 1476
Db 1485 GCACCAAGGCTCTGGGCCAGGCTGCGAGATGGGGCTGACCTGCTGGGGTCTTGGCGC 1544
QY 1477 ACACATGACACAGCTCTCCCACTCTGAGGCCGAGGAGTTCGTCTCCCGCTCTCCA 1536
Db 1545 CTGSCACATACCTTAGCGAATTCGCGCGGACACCCACTTCTCTGCACTGCGCAGCGCCA 1604
QY 1537 CCAGAAACTAC-----TTCGGCTCCCTGCGCGGAGCACAAGCAATGACTATGGGA 1590
Db 1605 GCTTCGGTTCACAGCAGCTCTTGGGGCTGCCCCGAGACCCAGGGAGCAGGCTCAGCGGCA 1664
QY 1591 CTTTCAACTTCTTCGGGGCGGCTGATGATCCCTAATACAGTATACAGCTCTCAGCTCTCATCC 1650
Db 1665 CTTTGGCTGCTGGTGGGAGCTCAGCATCCCGGCAAGGGGTGAGCTTGTCTGTGGTGC 1724
QY 1651 CCCAGATGCAATACCCCGAGGGAAGATCATGAGATCTACCTCAGCTGCACAAAGCCGG 1710
Db 1725 CCAATGGAGCAATTCGCCAGGCAAGTTCACAGATGATCTACTACTCATCAACAGGCGAG 1784
QY 1711 AAGACGTGAGGTTGCCCCCTAGCTGCTCAGACCCCTGCTGAGTCCCATGCTTAGCTGTG 1770
Db 1785 AAGTACCTGCGCTTTCAGAAAGGAGACCCAGACAGTATTTGAGCCCCCTCGGTGACCTGTG 1844
QY 1771 GACCCCTGCGCTCTGCTCAGCCGCGCAGTCACTCTGGCTATGGACCACTGTGGGGAGC 1830
Db 1845 GACCCACAGGCTCTCTGCTGTGCGGCCGCTCATCTCACCATGCCCCACTGTGCCAAG 1904
QY 1831 CAGCCCTGACAGCTGGAGCTGCGCTCAAAGACAGTCTGTCGAGGGCAGCTGGGAGC 1890
Db 1905 TCAGTGGCCGCTGACTGGATCTTTCAGCTCAAGACCCAGGCCACAGGGCCACTGGGAG- 1963
QY 1891 AGGATGTGCTGACCTTGGCGGAGGAGGCCCTCCACCTCTACTACTGCTGCTGAGGAGG 1950
Db 1964 --GAGTGTGACCTTGGATGAGGAGACCTGAACACACACCTGCTACTGCCAGTGGAGC 2021
QY 1951 CCAGTGCCTGCTACGCTCTTTCACCGGAGCAGTGGCGGCTTTCCCTGTGGGAGGAGGCC 2010
Db 2022 CCAGGGCTGTGCATCTCTGTGGACAGCTGGGCACTTACGTTTACGGGCGAGTCTCT 2081
QY 2011 TCAGGCTGCTGCCCGCAAGCCCTCAAGCTCTTCTGTGTCGCGGCTGGCTGCTGACCT 2070
Db 2082 ATTCCGCTCAGCAGTCAAGCGGCTCCAGTGGCGCTCTTCCGCCCGCCCTCTGCACT 2141
QY 2071 CCCTGAGTACAACTCCGGGTCTACTGCTGCTGATGACACCCAGATGCACTCAAGGAGG 2130
```



Db	2142	CCTGTGAGTACAGCCTCGGGTCTACTGCTGTGAGGACACGCGCTGTAGACTGAAGAGG	2201	APPLICANT: Guo, Xiaojia
Qy	2131	TGTTGAGCTGGAGAAAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGTCTCTGCACT	2190	APPLICANT: Miller, Charles E.
Db	2202	TGTTGAGCTGGAGCGAGCTCTGGCGGATCTTGGTGGAGGAGCGAAACCGCTAATGT		APPLICANT: Gangolli, Esha A.
Qy	2191	TCAAGGACAGTTACCAACACTGGCTTATCCATCCACGATGCCCCAGCTCCCTGTGGA		TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
Db	2262	TCAAGGACAGTTACCAACACTGGCTTATCCATCCACGATGCCCCAGCTCCCTGTGGA		FILE REFERENCE: 21402-214 CIP
Qy	2251	AGAGTAACTCTGTGCTGAGTACAGGAGATCCCTTTTATCATCTGGAATGGACGC		CURRENT APPLICATION NUMBER: US/10/087,684
Db	2322	GGAGCAAGCTGTGGCCAAATACAGGAGATCCCTTTTATCATCTGGAATGGACGC		CURRENT FILING DATE: 2003-03-10
Qy	2311	AGCGTACTTGTGCACTGACCTTACCTCTGGAGCGTGTACGCCCCAGCACTAGTACCTGG		PRIOR APPLICATION NUMBER: 60/253,834
Db	2382	AGAAGCCCTCCACTGCTGCTTACCTCTGGAGGACAGCTTGGCCCTCCACAGAGTCA		PRIOR FILING DATE: 2000-11-29
Qy	2371	CTGTCAAGCTGTGGTGTGGAGTGGAGGCGAGGCGAGAGTTCAGCATCAACTTCA		PRIOR APPLICATION NUMBER: 60/250,926
Db	2442	CCTGCAAGATCTGCTGCGCAAGTGAAGGGAGGCGGAGATATCCAGTGCATACCA		PRIOR FILING DATE: 2000-11-30
Qy	2431	ACATCACAAG--GACACAAGTTTCTGAGTCTGCTGCTGTGAGAGTGAAGCGGGG		PRIOR APPLICATION NUMBER: 60/264,180
Db	2502	CTGTGGAGAGACACTGTGCTGCTCCCTGGACACTCTCTGCTGCCCCCTGGCAGCTG		PRIOR FILING DATE: 2001-01-25
Qy	2488	TCCAGAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGAGTGGAGTAA		PRIOR APPLICATION NUMBER: 60/274,194
Db	2562	TCACCAAGCTGGAGCTTATGCTTCAAGATCCCACTGTCCATCGGCGAGAGATAT		PRIOR FILING DATE: 2001-03-08
Qy	2548	TTTCCAGCTGGAGCCACCTGTAGGCGGGTGGGCTGCGGAGTGGCGGACTGCGCCAGAAC		PRIOR APPLICATION NUMBER: 60/313,656
Db	2622	GCAACAGCTAGATGCCCCCAACTCACGGGCAATGACTGGCGGATGTAGCAGAGC		PRIOR FILING DATE: 2001-08-20
Qy	2608	TCCACCTGGAGAGCATCTCAGCTTCTTTGCTTCCAGCCAGCCAGCCACATGATCC		PRIOR APPLICATION NUMBER: 60/327,456
Db	2682	TCTCTATGGACCGTACTGATTAATTTGTCACCAAGAGCGAGCCAGGGTGTGATCC		PRIOR FILING DATE: 2001-10-05
Qy	2668	TCAACCTGTGGAGCGCGGCACTTCCCAACGGCAACTCAGCCAGCTGTGCTGAGCAG		NUMBER OF SEQ ID NOS: 220
Db	2742	TGGACCTCTGGAAGCTCTGCGAGGAGCAGTGGGACCTCAACAGCCTGGCGAGTGCCT		SOFTWARE: Curation version 0.1
Qy	2728	TGGTGTGAGTGGGCGAGCCAGCTGCGCTTTCACAGTGTGGAGCTGAGTGTGGA		SEQ ID NO 3
Db	2802	TGGAGGAGATGGCAAGAGTGAATGCTGTGGTGTGGTGTGGCCACCGAGGAGTGTGA		LENGTH: 2860
RESULT 14				TYPE: DNA
US-10-087-684-3				ORGANISM: Homo sapiens
; Sequence 3, Application US/10087684				FEATURE:
; Publication No. US20040029116A1				NAME/KEY: CDS
; GENERAL INFORMATION:				LOCATION: (59)..(2857)
; APPLICANT: Edinger, Shlomit R.				US-10-087-684-3
; APPLICANT: MacDougall, John R.				
; APPLICANT: Millet, Isabelle				
; APPLICANT: Ellerman, Karen				
; APPLICANT: Stone, David J.				
; APPLICANT: Grosse, William M.				
; APPLICANT: Lepley, Denise M.				
; APPLICANT: Rieger, Daniel K.				
; APPLICANT: Burges, Catherine E.				
; APPLICANT: Casman, Stacie, J.				
; APPLICANT: Spytek, Kimberly A.				
; APPLICANT: Boldog, Ferenc L.				
; APPLICANT: Li, Li				
; APPLICANT: Padigar, Muralidhara				
; APPLICANT: Mishra, Vishnu				
; APPLICANT: Shency, Suresh G.				
; APPLICANT: Rastelli, Luca				
; APPLICANT: Tchernev, Velizar T.				
; APPLICANT: Vernet, Corine A.M.				
; APPLICANT: Zerhusen, Bryan D.				
; APPLICANT: Malyankar, Uriel M.				

Qy	184	ACCAGTGCCTGTGTGTCACACCGGACCTGCTTCCCACTTCTTGGTGGAGCCCGAGT	243	Query Match 30.8%; Score 887.4; DB 13; Length 2860;
Db	168	ACTCTTCCCGTCAGCGCCAGCAGCGCTGCTTCTTCTCAGGAGCCACAGGACG		Best Local Similarity 60.6%; Pred. No. 5.4e-211;
Qy	244	TGTACATCGTCAAGAACAGCCAGTGTGTGTGCAAGCGCTGCCCCCGCAGCAGA		Matches 1636; Conservative 0; Mismatches 961; Indels 102; Gaps 7;
Db	228	CCTACATTTGTAAGAACAGCCCTGTGGAGCTTCTGCGCGCTTCCCCCGCAGCAGA		
Qy	304	TCTTCTTCAAGTCAACGGGAGTGGTGGCCAGCTGGACCACTGATCGAGCGCAGCA		
Db	288	TCTACTTCAAGTCAACGGGAGTGGTTCAGCCAGAACACCGTCAACAGGAAGGCC		
Qy	364	CAGACGGAGCAGTGTGAGCCGACCATGAGGTCCCGATTAAATGTCTCAAGGAGCAGG		
Db	348	TGGATGAGCCACCGGCTGCGGTCGCGAGGTGACATCGAGGTGTCGCGCAGCAGG		
Qy	424	TCGAGAGGTGTCGCGCTGAGGAATACTGTCAGTGGCTGGCATGGAGCTCCTCGG		
Db	408	TGGAGAGCTCTTGGGCTGGAGGATTAATGTCGTCAGTGGCTGGAGCTCCGAG		
Qy	484	GCACCAACAGTCAAGAGCTTACATCCCATAGCCAGATGTCGCAAGAACTTCGAGC		
Db	468	GCACCAACAGTCAAGAGCTTACATCCCATAGCCAGATGTCGCAAGAACTTCGATC		
Qy	544	AGGAGCCCTGGCCAGAGGTGTCCTTGGAGCAGGAGTCTGCTGCTGCTGCTGCTGCT		
Db	528	AGGAGCCCTGGCCAGAGGTGTCCTTGGAGCAGGAGTCTGCTGCTGCTGCTGCTGCT		
Qy	604	CGGAGGAGTCTCCAGCCAGGTGGAGTGGCTCCGGAACAGAGACCTGGTGGACCCGT		
Db	588	CGGAGGAGTCTCCAGCCAGGTGGAGTGGCTCCGGAACAGAGATGTCTCGACCCCA		
Qy	664	CCCTGGACCCCAAGTATATACATACCGGGAGCAGCGCTGGTGGTGGAGAGGCCGCC		
Db	648	CCGAGGACCAACACTTCTGCTTCCATCGACCAACCTCATCATCGCCAGGCCGCC		
Qy	724	TTGCTGACACGGCCAACTTACACTTGGTGGCCAAAGATCATGTGCGACGCTGCGCCG		
Db	708	TGTCGACACTGCCAACTATACCTCGTGGGCCAAGACATCTGTGCGCCAAACCGCG		





GENERAL INFORMATION:		US-10-218-779-3	
APPLICANT: Edinger, Shlomit		Query Match	30.8%; Score 887.4; DB 13; Length 2860;
APPLICANT: MacDougall, John		Best Local Similarity	60.6%; Pred. No. 5.4e-211;
APPLICANT: Millet, Isabelle		Matches 1636; Conservative	0; Mismatches 961; Indels 102; Gaps 7;
APPLICANT: Ellerman, Karen			
APPLICANT: Stone, David		Qy	184 ACCAGTCCCTGGTCCACCGGACCGGACCTGCTTCCCACTTCTGCTGGAGCCCGAGGATG 243
APPLICANT: Gerlach, Valerie		Db	168 ACTCTCCCGTACGCGCAGCAGGCGCTGCTGAGCTTCTTCTGAGAGCCACAGGAGC 227
APPLICANT: Grosse, William			
APPLICANT: Alsobrook II, John		Qy	244 TGTACATCGTCAAGAACCAAGCCAGTGTCTGTGTGCAAGCCCGTGCCTCCGACGACG 303
APPLICANT: Lepley, Denise		Db	228 CCTACATTTGTGAGAACCAAGCTGTGGAGCTTCTGCTGCGCGCTTCCCGCCACACAGA 287
APPLICANT: Rieger, Daniel		Qy	304 TCTTCTTCAAGTGCACCGGGAGTGGGTGCGGCGAGGTGACACGATCGAGTGTGCGGACGCA 363
APPLICANT: Burgess, Catherine		Db	288 TCTACTTCAAGTGCACCGGGAGTGGGTGAGCCAGACACACGCTCACAGGAAGGCC 347
APPLICANT: Casman, Stacie		Qy	364 CAGACGGAGCAGTGTGAGCCGACCATGGAGTCCGCAATTAATGTCTCAAGGACGACGAG 423
APPLICANT: Spytek, Kimberly		Db	348 TGGATGAGCCACCGGCTGCGGTGCGGAGGTGAGATCAGATGTCAGGTGTGCGGACGAGC 407
APPLICANT: Boldog, Ferenc		Qy	424 TCGAGAAGTGTTCGGGCTGGAGGAATACTGGTGGCAGTGCCTGGGATGAGCTCTCTGG 483
APPLICANT: Li, Li		Db	408 TGGAGGAGCTCTTTGGGCTGGAGGATTACTGGTGGCAGTGCCTGGGCTGGGCTCCGCGAG 467
APPLICANT: Padigaru, Muralidhara			
APPLICANT: Mishra, Vishnu			
APPLICANT: Patturajan, Meera			
APPLICANT: Shenoy, Suresh			
APPLICANT: Rastelli, Luca			
APPLICANT: Tchernev, Velizar			
APPLICANT: Vernet, Corine			
APPLICANT: Zerhusen, Bryan			
APPLICANT: Malyankar, Uriel			
APPLICANT: Guo, Xiaojia			
APPLICANT: Miller, Charles			
APPLICANT: Gangolli, Esba			
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
FILE REFERENCE: 21402-214			
CURRENT APPLICATION NUMBER: US/10/218,779			
PRIOR FILING DATE: 2002-08-14			
PRIOR APPLICATION NUMBER: 60/253,834			
PRIOR FILING DATE: 2000-11-29			
PRIOR APPLICATION NUMBER: 60/250,926			
PRIOR FILING DATE: 2000-11-30			
PRIOR APPLICATION NUMBER: 60/264,180			
PRIOR FILING DATE: 2001-01-25			
PRIOR APPLICATION NUMBER: 60/313,656			
PRIOR FILING DATE: 2001-08-20			
PRIOR APPLICATION NUMBER: 60/327,456			
PRIOR FILING DATE: 2001-10-05			
NUMBER OF SEQ ID NOS: 216			
SOFTWARE: Patent In Ver. 2.1			
SEQ ID NO 3			
TYPE: DNA			
LENGTH: 2860			
ORGANISM: Homo sapiens			
US-10-218-779-3			
Qy	484	GCACCAACAGAGTCAAGAGGCTCATCCGATAGCCAGATTGCGCAAGAACTTCGAGC	543
Db	468	GCACCAACAGAGTCAAGAGGCTCATCCGATAGCCAGATTGCGCAAGAACTTCGATC	527
Qy	544	AGGAGCCGCTGGCCAAAGGAGGTTCCTGGAGAGGGCATTCGTGCTGCGCTTCCCTGCGTCCAC	603
Db	528	AGGAGCCCTTGGGCAAGGAGGTGCGCCCTGGACCATGAGGTTCCTCTGCAATGCGCGCCG	587
Qy	604	CGAGGGCATCCCTCCAGCCGAGGTGGAGTCCGCGACGAGGACCTGGTGGACCCGT	663
Db	588	CGAGGGGGTGCCTGTGGCCGAGGTGGAAATGGCTCAAGAAATGAGGATGTCTATGACCCCA	647
Qy	664	CCCTGGACCCCAATGTATATACATCACCGCGGAGCACAGCCCTGGTGGTGGCGAGCCCGCC	723
Db	648	CCAGGACACCAACTTCCTGCTCACCATCGACCAACCTCATCATCGCCGACGCGCCGCC	707
Qy	724	TTGCTGACACGGCCAACTACCTCGGTGGCCAAAGAACATCGTGGGACAGTGGCGGAGCG	783
Db	708	TGTGGGACACTGGCCAACTATACCTGGTGGCCAAAGAACATCGTGGGCAAAACGCGGAGCA	767
Qy	784	CTTCCGCTGTCTCATCGTCTAGTGAACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	843
Db	768	CCATGCCACCGTCACTGCTTACGTGAATGGCGCTGGTCCAGCTGGGAGAGTGGTCCAC	827
Qy	844	TTGTCAGCGCCAGCTGGGCGGGGCTGGCAGAAACCGGAGCCGAGCTGCACCAACCCCG	903
Db	828	CTGTCTCAACCGCTGTGGCGGAGGTGGCAGAGCGCACCCGACCTGCACCAACCCCG	887
Qy	904	CGCCTCTCAAACGGGGCGCTTTCTGTAGGGGCGAATGTCTGATGAGCGGACCTGACCGACCTCTCT	963
Db	888	CTCCACTCAACGGAGGGGCTTCTGCGAGGGCCAGGCAATTCAGAA---GACCGCCTGCA	944
Qy	964	CTGTGCTGTCTGTGTGAGCGGACGTGGAGCCCGTGGAGCAAGTGTGGCGCTGTGGGC	1023
Db	948	CCACCATCTGCCAGTCTGATGGGCGGTGGAGCGGAGTGGAGCAAGTGTGACCTTCGAGCA	1004
Qy	1024	TGGACTGACCCCACTGGCGGAGCGGTGAGTGTCTGTGACCCAGCACCCCGCAACCGAGGGG	1083
Db	1005	CTGAGTGTGCCCACTGGCGGTGGGTGGTGTGACCGCGCCACCCAGCAACCGAGGCC	1064
Qy	1084	AGGAGTGCAGGGCACTGACCTGTGACACCCCGCAACTGTACAGTGTGCTGTGTACACA	1143
Db	1065	GTGACTGCAGCGGACGCTGTCTGACTTAAGAACTGCACAGATGGGTGTGCAATGCAAC	1124
Qy	1144	GTGCTTTCGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTGCTGCC	1203
Db	1125	TGGAGGCTTCAGGGGATGCGCGCGTGTATGCGGGGCTGCTGGTGGCCATCTTCGTGGTGG	1184
Qy	1204	TGGTCTGCTGCTGTGTCTCTCTCATCTCTGTTTATTGCGGAAAGAGAGGGGCTGGACT	1263
Db	1185	TGGCAATCCTCATGGCGGTGGGGTGGTGTGACCGCGCAACTGCGGTGACTTCGACA	1244
Qy	1264	CAGATGTGGTGTGCTGTTCATT---CTCAGCTAGGCTTCCAGCCCGCTCAGCATCAGC	1320
Db	1245	CAGACATCACTGACTCATCTGTGCTCTGACTGTGGTGGTTCACCCCGCTCAACTTAAGA	1304
Qy	1321	CCAGCAAGCAGACACCCCATCTGCT------CACCATCCAGCGGACCTCAGCACCA	1374
Db	1305	CGGCAAGCCGAGTAACCCGAGCTCTTACACCCCTCTGTGCTCTGACCTGACAGCA	1364
Qy	1375	CCACCACCTACAGGGCAGTCTCTGTCCCGGAGGATGGGCGCCAGCCCAAGTTCCAGC	1434
Db	1365	GGCGCGCATCTACCGCGGACCCGTGTATGCCCTGCTGAGGACTCCACCCGCAAAATCCCA	1424
Qy	1435	TCACCAATGGGCACTGTCTCAGCCCC-----	1460
Db	1425	TGACCAACTCTCTCTGCTGGACCCCTTACCCAGCCTTAAGTCAAGGTCTACAGCTCCA	1484
Qy	1461	-----	1476
Db	1485	GCACCAACGGGCTCTGGGCGACGCTGGCAGATGGGGCTGACCTGCTGGGGCTTGGCGC	1544



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 6, 2004, 11:46:57 ; Search time 7179 Seconds

(without alignments)

11983.973 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccgggtgag.....cctccccacacggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gsl1:\*

29: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907.8	31.5	2802	29	AY406491 Homo sapi
2	898.6	31.2	2791	29	AY406493 Mus muscu
3	871.6	30.3	3790	11	AK031655 Mus muscu
4	818.4	28.4	3866	11	AK018177 Mus muscu

5	763.2	26.5	2532	29	AY411747	AY411747 Homo sapi
6	763.2	26.5	2802	29	AY406492	AY406492 Pan trogl
7	756.2	26.2	1034	12	BI758231	BI758231 603029876
8	726.4	25.2	2532	29	AY411749	AY411749 Mus muscu
9	726.2	25.2	909	13	EX364574	EX364574 BX364574
10	722.2	25.1	788	9	AI951556	AI951556 wv36f04.x
11	711.2	24.7	1532	11	BC033727	BC033727 Homo sapi
12	710	24.6	818	12	BI818609	BI818609 603033362
13	709.8	24.6	843	13	EX348193	EX348193 BX348193
14	700.2	24.3	939	13	EX345406	EX345406 BX345406
15	679.4	23.6	1040	13	EX345407	EX345407 BX345407
16	664.2	23.1	1201	13	EX422753	EX422753 BX422753
17	664	23.0	889	13	BQ691915	BQ691915 AGENCOURT
18	653.6	22.7	934	10	BF311804	BF311804 601897316
19	648.6	22.5	2532	29	AY411748	AY411748 Pan trogl
20	646.4	22.4	859	10	BF311896	BF311896 601897733
21	644.2	22.4	874	13	BQ689148	BQ689148 AGENCOURT
22	604.4	21.0	755	13	BU612387	BU612387 UI-M-EWO-
23	596.4	20.7	662	9	AL516580	AL516580 UI-M-EWO-
24	579	20.1	2775	29	AY401471	AY401471 Mus muscu
25	573.4	19.9	604	10	BE314370	BE314370 601147261
26	573.4	19.9	2775	29	AY401469	AY401469 Homo sapi
27	554.6	19.3	1175	10	BF530640	BF530640 602071931
28	548	19.0	788	14	CA317532	CA317532 UI-M-FW0-
29	539.4	18.7	678	14	CA749784	CA749784 UI-M-FW0-
30	524.4	18.2	824	12	BI737024	BI737024 603360874
31	493	17.1	1030	12	BM479478	BM479478 AGENCOURT
32	471.2	16.4	540	12	BI467799	BI467799 389583 MA
33	467.8	16.2	917	13	BX453456	BX453456 BX453456
34	463.6	16.1	1028	13	BX331575	BX331575 BX331575
35	449.8	15.6	987	13	BX452510	BX452510 BX452510
36	445	15.4	2507	29	AY401470	AY401470 Pan trogl
37	433	15.0	471	13	BX282095	BX282095 BX282095
38	430.6	14.9	675	14	CA315487	CA315487 UI-M-FW0-
39	424.4	14.7	843	12	BG913440	BG913440 602811321
40	424	14.7	780	12	BG470852	BG470852 602511705
41	416	14.4	718	14	CA318180	CA318180 UI-M-FW0-
42	415.4	14.4	1147	13	BU840446	BU840446 AGENCOURT
43	413	14.3	960	12	BG329812	BG329812 602429458
44	411.6	14.3	1111	12	BG298307	BG298307 602337080
45	411.4	14.3	749	14	CF735417	CF735417 UI-M-HB0-

## ALIGNMENTS

RESULT 1	AY406491	AY406491	2802 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	Homo sapiens	Homo sapiens	UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,			
DEFINITION	AY406491	AY406491	Genomic survey sequence.			
ACCESSION	AY406491	AY406491	GI:39762465			
VERSION	AY406491.1	AY406491.1	GI:39762465			
KEYWORDS	GSS.	GSS.				
SOURCE	Homo sapiens	Homo sapiens	(human)			
ORGANISM	Homo sapiens	Homo sapiens				
REFERENCE	1	(bases 1 to 2802)				
AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.					
TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios					
JOURNAL	Science 302 (5652), 1960-1963 (2003)					
PUBMED	14671302					
REFERENCE	2	(bases 1 to 2802)				
AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.					
TITLE	Direct Submission					

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source Location/Qualifiers

1..2802

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

<1..>2802

/gene="UNC5C"

/locus\_tag="HCM2575"

ORIGIN

Query Match 31.5%; Score 907.8; DB 29; Length 2802;

Best Local Similarity 60.3%; Pred. No. 3.9e-142;

Matches 1660; Conservative 0; Mismatches 1014; Indels 78; Gaps 7;

98 GCCCGGCGCTGGCCAGCGCTCTCTGGGCATAGTCTCGCGCTTGGCTCGCGGCTCGGG 157

66 GCTCGTCTACCTGCGCTGCTCTCAGCGCGCGGCACTGGCTCCGCGCCCAAGA 125

158 TGCCCGCAGAGTCCACCGTGGCCAAACCGAGTCCCTGGTGGCAACCGGACCTGCTTCC 217

126 TGATGACTTTTTCATGAATCTCCAGAACTTTTCTCTGTATCCACTGAGCTCTGCC 185

218 CCATTCTCTGTGGAGCCCGAGTGTGTATCGTCAAGAAACAGCCAGTGTGCTTGT 277

186 ACATTCTCTTATTTAGCCTGAAGAAGCTTATTTGTGAAGAAATGAAGCTGTGAACCTGTA 245

278 GTGCAAGCGCGTGGCCGACGAGATCTTCTCAAGTGCACCGGAGTGGTGGCCCA 337

246 CTGTAAAGCAAGCCCTGCGCCAGATCTATTTCAAGTGTAAATAGTGAATGGGTTCATCA 305

338 GGTGGACACGTGATCGAGCCGACGACGCGGAGCGAGTGTGGTGGCCGACCATGAGGT 397

306 GAAGGACACATATGATGATGAAGAGTATGAACTTCCGCTCTCATTTGTCGGGAGT 365

398 CCGCATTAATGTCTCAAGGACGAGTGTGGAAGGTGTTCGGTGGAGGAATATCTGGTG 457

366 GAGCATTCAGATTTCGCGCCAGCAAGTGAAGAACTCTTTGGAAGTGAATATCTGGTG 425

458 CCAGTGGCGTGGATGAGCTCTCGGACACACGAGAGTCAAGAGGCTCATCCGAT 517

426 CCAGTGTGTGCTGGAGTCTCGCGGTACCAAGAGCGCGGAGCGGTATGTGCGGT 485

518 AGC-----CAGATTGCGCAAGACTTCGAGCAGGAGCGGCTGGCCAAAGGAGGTGTCCCT 571

486 TGCAATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 545

572 GGAGCAGGCGCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631

546 GGAACAGGAAGTCTTACTTCCAGTGTGACACCACTGAAGGATCCCAAGTGGCTGAGGTGA 605

632 GTGGCTCCGACAGGAGGAGTGTGAGACCCCTGCTGAGACCCCAATGATATCATCACGCG 691

606 ATGTTGTAATAATGAAGACATAATATGATCCCGTTGAAGATCGGAATTTTATATTAAT 665

692 GGAGCAGCAGCTGGTGGTGGCGAGCGCGCTGCTGACAGCGGCAACTACACCTCGGT 751

666 TGATCACAACCTCATCAATAAGCAGGCGCGATCTCTGATCTGCAATATACACCTGTGT 725

752 GCCCAAGAACATCGTGGCAGCTGCGCGACGCGCTCGGTGTGTGTGTGTGTGTGTGTGTGNA 811

726 TGCCAAAACATTTGTTCCAAAGAGGAAAGTACACTGGCACTGTGATAGTCTATGTCAA 785

812 CGTGGTGGTGTGAGTGAACCGAGTGTCCGTCTGACGCGGACAGTGTGGCGGGGTG 871

786 CGGTGGCTGTGCTACCTGAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA 845

872 GCAGAAACGAGCGGAGTGTGACCAACCGCGGCGCTCTCAACGGGGGGCGCTTCTGTGA 931

846 TCAGAAACGTACAGGACTTGTACCAACCGCGGACCACTCAATGGGGGTGCCTTCTGTGA 905

932 GGCGCAGATGTCCATGACCGCACCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 991

906 AGGGCAGATGTGCGAGAA---AATAGCTGTACTAGCTTATGCCAGTGGATGGCAGGTG 962

992 GAGCCCGTGGAGCAAGTGTGCGGCTGTGGGCTGGAGCTGACCCACTTGGCGGAGCCGTGA 1051

963 GACGCCATGGAGCAAGTGTGTACTTGTGGAATGAGTGCACCCACTTGGCGGAGGAGGA 1022

1052 GTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCAGGGGACACTGACCTGGACAC 1111

1023 GTGCACGGCGCCAGCCCCCAAGAATGGAGCAAGGACTGGAGCGGCTGTCTTGTGCATC 1082

1112 CGCAACTGTACCAGTGACCTCTGTGTACACAGTGTCTTGGGCCCTGAGGACGTGGCCCT 1171

1083 CAAGAACTGCATGATGGGCTTTGCATGCAGACTGTCTCTGATTCAGATGATGTGCTCT 1142

1172 CTATGTGGG---CCTCATGCGCGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1228

1143 CTATGTGGGATTTGATAGCAGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1202

1229 CTTCTGTTATTGCGGAAGAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATCTCT 1288

1203 GTTTGTGTATCGGAAGAATCATCTGTGACTTTGAGTTCAGATTTATTTGACTCTTCGGCACT 1262

1289 CACTCAGGCTTCCAGCCCTGACGATCAAGCCCGACGAAAGACGACAAACCCCATCTGCT 1348

1263 CAATGGGGGCTTTCAGCTGTGACATCAAG-----GCAGCAAGACAAAGATCTGCT 1313

1349 CACCATCCAGCCGAGCTCTGACGACCAACCACTACAGGGGAGTCTCTGTCTCCCGCA 1408

1314 GGCTGTATCCCGCAGACCTCTACGTCAGCTGACGACCATGTACAGGAGACTCTGTATGCGCT 1373

1409 GGATGGCCCGAGCCCGAGTTCAGCTCAGCAATGGGCACTGTCTGAGCCCTCTGGGTG 1468

1374 GCATGACGTCTCAGCAAAATCCCAATGACCACTCTCCAAATCTGGATCCACTGCCCAA 1433

1469 CGCGCGCACACACTGACCAACACAGCT-----CTCCACCTCTGAGGCGGAGGA 1516

1434 CCTGAAATCAAGTGTACACACCTCAGGTGTGTGACGCCCCCAAGATGACCTCTCTGA 1493

1517 GTTCGTCTCCGCTCTCTCCAGCAAACTCTTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1576

1494 GTTTACGTCCAAAGTGTCTCCCTCAGATGACCCAGTCTGTGTGGAGAAATGAAGCCCTCAG 1553

1577 CATGA-----CCTATGGAGACCTT 1594

1554 CCTGAAGAACAGAGTGTAGCAAGGAGACTGATCCATCTCTGATCCGCAATTTGGCAGCTT 1613

1595 CAACCTTCTCCGGGCGCGGCTGATCCCTTAATAACAGGTATCAGCTCTCTCATCCCCC 1654

1614 CAATCTGCTGGAGGTCACTTTATTTCCCAATTCAGGAGTTCAGTGTGCTGATTCGCGC 1673

1655 AGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCTGACAGGCGGAGGA 1714

1674 TGGGGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACAGGAAGAAGAAC 1733

1715 GTGAGGTGCCCCCTAGCTGTGTCAGACCTGTGAGTCCCATCTGTAGCTGTGGAC 1774

1734 TATGAGGCCACCCATGAGTACTCTCAGACACTTTTGGACCTTTGGTGTGAGCTGTGGCC 1793

1775 CCTGGCGCTCTGTCTCACCCGCGGCAAGTCTCTGTGATGACCACTGTGGGAGCGCCAG 1834

1794 CCCAGAGCTCTGCTCACCGCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1853

1835 CCTGACAGCTGGAGCTCGGCTCAAAAAGCAGTCTGCGGAGGCGAGCTGGGAGCAGGA 1894

1854 TACCGAGGACTGGAATACTGCTCAAGAACCCGAGCAGCAGGAGCAGTGGG---AGGA 1910

1895 TGTGTGTGACCTGTGGCGAGGAGCGCTCTCCACCTCTACTACTGTGACCTGTGAGGCCAG 1954

1911 TGTGTGTGTGTGGGAGGAGAACTTTCACACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1970

```
QY 1955 TGCTGTCTACGCTTTCACGAGAGCTGGCGCTTTTGCCTGTGTGGAGAGCGCCTCAG 2014
Db 1971 GGCCTGCGACATCTCTCAGAGAACTCAGACACCTCAGACCTAGCCCTGTAGACATTCACAC 2030
QY 2015 CGTGGCTCCCGCCAGCGCTCAAGCTCTTCCTGTGTGGCGGGTGGCTGACCTCCCT 2074
Db 2031 CAAGCGGCTCGAAGCGCCCTCAAGCTGGCCATCTTTGGGCGCCTGTGCTGCTCTGCT 2090
QY 2075 CGAGTACAACTCCGGGTCTACTGCTGCTGATGACACCCAGCATGCTCAAGAGGGTGT 2134
Db 2091 GGAGTACAGATCCGAGTCTACTGCTGATGACACCCAGCATGCCCTGAAGAAATTTT 2150
QY 2135 GCAGCTGGAGAGAGCTGGGGGAGAGCTGATCCAGAGAGCCAGGGTCTTGCACTTCAA 2194
Db 2151 ACATCTTTGAGAGACAGATGGGAGGACAGCTCCTAGAGAACTTAAGGCTCTCTATTAA 2210
QY 2195 GGAGAGTTTACCACACCTGGCCCTTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGAG 2254
Db 2211 AGGAGAGCCACCAACCTGGCGCTGTCAATTCAGATATGCCCATTTCCCTCTGGAAGAG 2270
QY 2255 TAAGCTCTTTGTAGCTTACAGGAGATCCCTTTTATCAGATCTGGAATGCGAGGAG 2314
Db 2271 CAAATCTGCTGCTAAATATCAGGAATTCATTTTACATGTTTGGAGTGGATCTCAAG 2330
QY 2315 GTACTGTGCTGACCTTCACTCCCTGGAGCGTGTGAGCCCGAGCACTAGTACCTGGCGTG 2374
Db 2331 AAACCTGTGCTGACCTTCACTCTGGAAGATTTAGCTTGAACACAGTGGAGCTGTTTG 2390
QY 2375 CAAGCTGTGGTGTGGAGCTGGAGGCGAGGCGAGAGCTTCCAGTCACTTCAACAT 2434
Db 2391 CAAACTCTGTGTGGCGAGGTGGAAGAGAGAGGCGAGATCTCCAGCTCAACTGCACCGT 2450
QY 2435 CACCAAGGACACAAGGTTTCTGAGCTGTCTGGAGCTGTGAGAGTGAAGCGGGTCCAGC 2494
Db 2451 GTCAGAGGAACCTACTTGGCATCATTTGCGCTGTGGATCTCGGAAACACCATCACCAC 2510
QY 2495 CTGTGTGGGCGCCAGTGTCTTCAAGATCCCTTTCTCAATTCGGCAGAGATAATTTCCAG 2554
Db 2511 GGTACGCGGGCGCCAGTGTCTTCAAGATCCCTTCTCCCTATCCGCGAGAGCTCTGTAGCAG 2570
QY 2555 CTGGAGACCACTGTAGGCGGGTGCAGTCTGGCGGACTTGGCGGACTTGGCGGAGAACTCCACCT 2614
Db 2571 CTGGATGCCCCCGACAGAGAGGCGCATGACTGGAGGATGCTGGCCCATAAAGCTGAACCT 2630
QY 2615 GGACAGCATCTCAGCTTCTTTCCTTCCAGCCAGCCAGCCAGCCAGCATGATCTCAACCT 2674
Db 2631 GGACAGTACTTGAATTACTTTGCCACCAATCCAGCCCAACTGGCGTAACTCTGATCT 2690
QY 2675 GTGGAGGCGGCGCACTTCCCAAGCGCACTCAGCCAGCTGGCTGCGAGAGTGGCTGG 2734
Db 2691 TTGGGAAGCACAGAACTTCCAGATGGAACCTGAGCATGCTGGCAGCTGCTTGAAGA 2750
QY 2735 ACTGGGCCAGCAGAGCTGGCTCTTTCAGTGTGCGAGGCTGAGTGTGA 2786
Db 2751 AATGGAGACATGAACGGTGTGTCTTACGAGCAGAGAGGCGAGTATTA 2802

RESULT 2
AY406493
LOCUS
DEFINITION
AY406493 Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
VERSION
AY406493.1 GI:39762467
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2791)
AUTHORS
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
```

```
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2791)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. 2791
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2791
/genes="UNC5C"
/locus_tag="HMC2575"
ORIGIN
Query Match 31.2%; Score 898.6; DB 29; Length 2791;
Best Local Similarity 61.5%; Pred. No. 1.4e-140;
Matches 1635; Conservative 0; Mismatches 939; Indels 83; Gaps 9;
QY 190 TGCTGTGGTCCCAACCGGACCTGCTCTTGTGTGAAGCGCGTGTGGAGCGCGAGGATGTGTACA 249
Db 158 TTCCATCTGACCCACCTGAGCCATTGCCACATTCCTCATTTGAGCCCGAGAAAGCTTACA 217
QY 250 TCGTCAAGAAACAGCCAGTGTCTTGTGTGAAGCGCGTGTGGAGCGCGGATCTTCT 309
Db 218 TTGTGAAGAAACAGCCCTGTGAACCTGATTTGAAGCGGAGCCCTGCGCCAGATCTACT 277
QY 310 TCAAGTGCACAGGGAGTGGGTGGCGCCAGTGCACACCGTGTGAGCGCGAGCAGACAG 369
Db 278 TCAAGTGCACAGCGAGTGGGTTCATCAGAAGGACCACTAGTAGAGAGATAGATG 337
QY 370 GGAGCAGTGGTGAAGCGGACCATGAGGTCGGCTCGGCTTAATGTCTCAAGCGAGAGTGCAGA 429
Db 338 AAACCTCTGCTTAATTTGTGAGAGAGTGAAGTTCAGATTTCCAGCCAGAGTGGAGG 397
QY 430 AGGTGTTCCGGCTGGAGGAATACTGTGCGCAGTGTGCGATGCGAGCTCTCGGGCACCA 489
Db 398 AACTGTTTGGGCTGAAGATTACTGTGCGCAGTGTGTGCGCTGGAGCTCAGAGGCACTA 457
QY 490 CCAAGAGTCAGAAAGGCTATCCCGCATAGCCAG-----ATTGCGCAAGAACTTCGAGC 543
Db 458 CGAAGAGTCGGAAGGCATACGTGCGCATTCGCTTTCCAGATCTGCGGAAGACATTCGAGC 517
QY 544 AGGAGCGCTGGCCAGGAGTGTCCCTGGAGCGGCGATCGTGTGCGCTGCGCTGCGCTCCAC 603
Db 518 AGGAACCTTTGGGAAGGAAGTGTCTTGGAGCAGGAAGTCTTACTCCAGTGTGCGCCAC 577
QY 604 CGAGGGCATCTCCCTCCAGCGGAGTGGAGTGTCTCCGGAACGAGGACCTGGTGGACCCCT 663
Db 578 CTGAAGGAGTCC-----GTAGATAGCTAAAGATGAAGACATAATTTGATCTCG 626
QY 664 CCTCGAGCCCAATGTATACATCAGCGGGAGCAGAGCTGGTGTGGCAGACGCGCCGCG 723
Db 627 CTGAAGATCGGAACCTTTTATATTACTATCATCAACCTGATCATCAAGCAAGCCCGAC 686
QY 724 TTGCTGACAGCGCCCACTACACCTGCTGGCGCAAGAAATCGTGGCAGCTGCCCGAGCG 783
Db 687 TCTCAGATACAGCAAAATATACCTGTGTGCCAAAATATTTGTTGCCAAGAGAAAAAGCA 746
QY 784 CCTCGCTGCTGTCTATCGTCTACGTGAAGCGTGGGTGTGCGAGCTGCGACCGAGTGTCCG 843
Db 747 CCACAGCCACTGTCTATCGTGTATGTTAATGTGTGCTGTCCACCTGGAGACAGAGTGTCTG 806
```

844 TCTGAGCCGACGCTGTGGCGCGCTGGCAGAAACGAGCCGAGCTGCACCAACCCG 903  
 807 TGTGTAAACAGCCGCTGTGGCGCAGGATATCAGAAACGCAAGAACTGCACCAACCCAG 866  
 904 CGCCTCTCAACGGGGCGCTTCTGTGAGGGGAGAAATGCCATGACCCGACCCGCTCTCT 963  
 867 CCCCACTCAATGTGGGGCTTCTGTGAGGGGAGAGTGTGCAGAA---AATAGCATGCA 923  
 964 CTCTGCTGTCTCTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGGTGGCGCTGTGGGC 1023  
 924 CTAGGTTATGTCAGTGGATGTAGGTGGACTTTCATGGAGCAAAATGGTCAACCTGTGGGA 983  
 1024 TGAGCTGACCCACTGGCGAGCCGTCAGTGTCTGACCCAGCAGCCCGCAACGGAGGGG 1083  
 984 CTGAATGCAACCCACTGGCGAGGAGGAGTGTACAGCAGCCAGCCCGCAAGAACGGGGTA 1043  
 1084 AGAGTGCACAGGCACTGACCTGCACACCCGCAACTGTACAGTACCTCTGTGTACACA 1143  
 1044 AGACTGTGATGGCTGTCTCCATCCAGAACTGCATGATGGCTGTGCATGCAGG 1103  
 1144 GTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGG---CCTCATCGCGTGGCCGTCT 1200  
 1104 CTGCTCTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAAACGTCT 1163  
 1201 GCCTGTCTGCTGCTGTCTCTCATCTCTGTTATTGCCGGAAGAGGGGGCTGG 1260  
 1164 GTCTGGCGATCACTGTGTGGTGGCCCTGTGTGTGATCGGAAGAACCCCGTACTTTG 1223  
 1261 ACTCAGATGTGGTGTACTGTGCTCACTTCTCAGCTCAGGCTTCCAGCCCGTCAGCATCAAGC 1320  
 1224 AGTCTGATCATATGATCTCTCAGCACTCAATGGCGCTTTCAGCTGTGAATCAAGG 1283  
 1321 CCAGCAAGCAGACAAACCCCATCTGTCTACATCCAGCGGAGCTCAGACACCAACCA 1380  
 1284 CTGCCAGACAGATCTCTGTGCTGTCCCGCTGACC-----TCACCTCAGCTGCAGCCA 1337  
 1381 CCTACAGGCGAGTCTGTGCTCCCGCAGGATG-----GGCCAGCCCAAGTTCCAGCT 1435  
 1338 TGTACAGGGGACCTGTCTATGCTGTGATGATGTCTCAGACAAATCCCATGACCAACT 1397  
 1436 CACCAATGGGCACCTGTCTACGCCCTGGTGGGGCGCCGACACATCTGCACCAACAGCTC 1495  
 1398 CTCGAATCTGGACCCACTACCAACTTGAATCAAGTGTACACAGCTCAGTGTG 1457  
 1496 TCCACCTC-----TGAGGCGAGAGTGTGCTCCCGCTCTCCACCC----- 1539  
 1458 TCACTCTCAGGATGACCTTGCAGGATCTCTATCCAACTGTCAACCCAGATGACCCAGT 1517  
 1540 -----AGAACTACTTCCGCTCCCTGCCCCGAGGCA 1569  
 1518 CTTGTGAGAAATGAGGCCCTTAACTGTGAAGAACAGAGCTTGCAGACAGACTGACC 1577  
 1570 CCAGCAACATGACCTATGGGACCTTCACTTCTCGGGGGCGGCTGATGATCCCTAATA 1629  
 1578 CATCTGCACAGCATTTGGTACCTTCACTCTCTTGGGGGTCACTCATCATTTCTTAAT 1637  
 1630 CAGTATCAGCTCTCTATCCCCCGAGATGCGATACCCCGAGGAGAGATCTATGAGATCT 1689  
 1638 CAGGATGAAGCTTGTGATTCCTCGGTGGGGCAATCTCTCAGGGGAGAGTCTATGAATGT 1697  
 1690 ACCTCAGCTGCACAGCCGAGACGTGAGTGTCCCTAGCTGTGTCAGACCTGTC 1749  
 1698 ATGACTGTACACAGGAAGAAATATGAGGCCCCCATGGAAGACTCTCAGACCTTAC 1757  
 1750 TGAGTCCCATCGTTAGCTGTGAGCCCTCGGTGCTCTCTCACCCTGGCCAGTCTATCTGG 1809  
 1758 TTAACCTGTGTGAGCTGTGGCTCTCTGAGGCTCTCTGACCCGCTGTCTATCTCA 1817  
 1810 CTATGACACACTGTGGGAGCCAGCCCTGACAGCTGAGCTGGCTGGCTCAAAAGAGT 1869  
 1818 CTCTGATCACTGTGAGACCCAGCAGCCAGGAGTGTGAAGATCAGCTCAAAACAGG 1877  
 1870 CGTGGAGGGGACCTGGGAGCAGGATGTCTGACACTGTGGCGGAGGAGCGCCCTCCAC 1929

1878 CAGTCAGGAGCAATGGGAGGATGTGTGGTG---GTTGGGAGGAGAACTTTCACAAACC 1934  
 1930 TCTACTATGCGAGCTGGAGGCGAGTGTCTACGCTTCTTCCCGAGAGCTGGGCGGCT 1989  
 1935 CTTGTTTACATTCAGCTGGATGACAGGCTTGGCATATCTCTCAGAGAACCTCAGTACCT 1994  
 1990 TTGCCCTGTGGGAGAGGCCCTCAGCGTGGCTGCCCGCAAGCCCTCAAGCTCTCTGT 2049  
 1995 ATGCCCTGTGGGAGTCCACCAACAGCAGCTGCCAAGCGCTTAAACTGSCCATCT 2054  
 2050 TTGCGCCGTGGGCTCCACCTCCCTCGAGTACAAATCCGGGTCTACTGCTCGCATGACA 2109  
 2055 TTGGGCCCCCTGCTCTCTTCCCTGGAGTACAGCATTAAGTCTACTTGCCTGGATGACA 2114  
 2110 CCACATGACTCAAGGAGGTGGTGCAGCTGAGAGCAGCTGGGGGACACGCTGATCC 2169  
 2115 CACAGATGCCCTGAAAGAAATTTCTAAGCTGGAGAGGCAATGGGAGGACAGCTCTAG 2174  
 2170 AGGAGCAGCGGCTCTGCACTTCAAGGACAGATTACCAACACCTGCGCCTATCCATCCAG 2229  
 2175 AAGAACCCAGGCTCTTCAATTTAAAGGAGCATCCACACCTGCGCTCTCTATTATG 2234  
 2230 ATGTGCCAGCTCCCTGTGGAAGAGTAAGTCTCTGTGACGTACACAGGAGATCCCTTTT 2289  
 2235 ACATGCCCATTCCTCTGGAAGAGCAAAATGTGCTGAAGTATCAGAAATTCATTTT 2294  
 2290 ATCAGATCTGAATGACGCGGCTTCTTGAAGTGTGCACTGCACCTTCAACCTCGAGCGTCA 2349  
 2295 ACCACATCTGAGTGGCTCTCAAGAAACCTCCAGCTGCACCTTCACTCTGGAAGACTCA 2354  
 2350 GCCCAGCACTAGTACCTGCGCTGCAAGTGTGGGTGTGGCAGGTGGAGGGGACGCGG 2409  
 2355 GCCTAACAACAGTGGAACTGTTTGCAAACTCTGTGTGGCGAGTTTGAAGAGAGGGC 2414  
 2410 AGAGCTTTCAGCATCAACTTCAACATCAACAGGACCAAGTTTGTGAGCTGTGCTGCTC 2469  
 2415 AGATCTTCCAGCTCAACTGTACTGTGTGAGAGGAACTTCTGCGATCGACTTACCTCTCC 2474  
 2470 TGGAGATGAAGCGGGGCTCCAGCCCTGTGGCGCCCGAGTCCCTTCAAGATCCCTTCC 2529  
 2475 TGGACCTGTGTAGTACCATCAGCTGTCCAGGACCAAGTGTCTTTCAGCATCTCTCC 2534  
 2530 TCATTGGCAGAGATAATTTCCAGCTGGAACCCACCTCTGTAGCGGGGTGCGAGCTGGC 2589  
 2535 CTATCCGCGAGAGCTATGAGCAGCTGTGATGCCCTCAACAGAGGCCATGACTGGA 2594  
 2590 GGACTCTGGCCCGAAGAACTCCACTGGAACAGCCATCTCAGCTTCTTTGCTTCCAGCCCA 2649  
 2595 GGATGCTGGCCCAATAACTCAACCTGGACAGTACTTGAATTAATTTGGCCCAAAATCGA 2654  
 2650 GCGCCACAGCATCATCTCAACCTGTGGAGGCGCGCACTTCCCAACGGCAACCTCA 2709  
 2655 GCGCACTGGCGTAACTCTGGATCTTTGGGAAGCAGAGACTTCCAGATGGAAACCTGA 2714  
 2710 GCCAGCTGGCTGACAGCTGGCTGGAGCTGGGCCAGCCAGAGCTGGGCTCTTTCAGTGT 2769  
 2715 GCATGCTGGCAGCGCTCTTGAAGAAATGGAAGACATGAGACAGTGTGCTTGGCAG 2774  
 2770 CGGAGGCTGAGTCTGA 2786  
 2775 CAGAAGGACAGTATTGA 2791

RESULT 3  
 AK031655

LOCUS  
 DEFINITION

AK031655 3790 bp mRNA linear HTC 18-SEP-2003  
 Mus musculus 13 days embryo male testis cdna, RIKEN full-length  
 enriched library, clone:6030473H24 product:unc5 homolog (C.  
 elegans) 3, full insert sequence.

ACCESSION  
 VERSION  
 KEYWORDS

AK031655  
 AK031655.1 GI:26327502  
 HTC; CAP trapper.



**SOURCE**  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**  
AUTHORS  
Carninci, P. and Hayashizaki, Y.  
TITLE  
High-efficiency full-length cDNA cloning  
JOURNAL  
Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE  
99279253  
PUBMED

**REFERENCE**  
AUTHORS  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL  
Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE  
20499374  
PUBMED  
11042159

**REFERENCE**  
AUTHORS  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
JOURNAL  
Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE  
20530913  
PUBMED  
11076861

**REFERENCE**  
AUTHORS  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE  
Functional annotation of a full-length mouse cDNA collection  
JOURNAL  
Nature 409, 685-690 (2001)  
MEDLINE  
11076861

**REFERENCE**  
AUTHORS  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
TITLE  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
JOURNAL  
Nature 420, 563-573 (2002)  
MEDLINE  
12011111  
PUBMED  
12011111

**REFERENCE**  
AUTHORS  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saito, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

**COMMENT**  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. 3790

**FEATURES**  
source

QY 670 ACCCAATATATACATCAACGCGGAGCAAGAGCTGTGTGTGCGAAGGCGCGCTTGCTG 729  
Db 755 ATCGGAATTTTATATTAATCATGATCAACAACCTGATCATCAAGCAAGCCGACTCTCAG 814  
QY 730 ACAGCGCAACTACACTGCTGGCGGCAAGCAACATCGTGACAGTGCAGCGCGCTCCG 789  
Db 815 ATACAGCAATATATACCTGTGTGCGCAAAATATTTGTCGAAGAAAGAACACACAG 874  
QY 790 CTGCTCTCATGCTACGCTGAACGCTGGGTGTGACAGTGGAGCCGAGTGTGCTGTGCA 849  
Db 875 CCACTGTCTATGCTGTATTTAATGTGTGCTGGTCCACCTGGACAGAGTGTGTGTGTA 934  
QY 850 GCGCCAGCTGTGGCGCGGCTGGGAGAAACGAGAGCGGAGCTGCACCAACCGCGGCTC 909  
Db 935 ACAGCGCTGTGGCGGAGGATATCAGAAACGCAAGAACCTGCACCAACCGAGCCGAC 994  
QY 910 TCAACGGGCGCTTTCTGTGAGGGGAGAGTGTCCATGACCCAGCCGCTCTCTCTGCG 969  
Db 995 TCAATGGTGGGGCTTCTGTGAGGGGAGAGTGTGCAGAA---AATAGCATGCACTAGT 1051  
QY 970 TTGTCTCTGTGGAGCGGAGCTGGAGCCCGTGGAGCAAGTGTGTGGGCTGTGGGCTGACT 1029  
Db 1052 TATGTCCAGTGGATGTGTAGTGTGACTTTCATGGAGCAATGTCAACCTGTGGGACTGAAT 1111  
QY 1030 GCACCCACTGTGGAGCGGCTGTGCTGTGACCCAGCACCCCGCAAGCGAGGGGAGGT 1089  
Db 1112 GCACCCACTGTGGAGCGGAGGTGTACAGCACCGCCCGCAAGACGGGGGTAAAGACT 1171  
QY 1090 GCCAGGCACTGTACTGGACACCGCAACTGTACAGTACCTCTGTGTACACAG-----TGCTTCTG 1144  
Db 1172 GTGATGGCCTGTCTCCATTCGAATCCAGAACTGTGACTGTGAGGGCTGTGATCGAGGATTC 1231  
QY 1145 -----TGCTTCTG 1152  
Db 1232 TTTACCCCAATTCACCTGAGCACAGACCCAGCAATGAATATGGATTTTCTTCTGCTCTG 1291  
QY 1153 GCGCTGAGGAGTGGCGCTCTATGTGGG---CCTCATGCGGTGGCGCTGTGCTGTGCTGCC 1209  
Db 1292 ACTCAGATGATGTGGCTCTCTACGTGGGATGTGTGATCGAGAACCCCGTGTGCTGGCGA 1351  
QY 1210 TGCTGCTGTGTCTCATCTCTGTTTATTCCGGAAGAGGAGGGGCTGACTCAGATG 1269  
Db 1352 TCATGTGTGTGGCGCTCTGTGTGTATCGAAGAACCCCGTGTGACTTTGACTGTGACA 1411  
QY 1270 TGGCTGACTGTCTCACTCTCAGCTTCCAGCGGCTGAGCATCAAGCCCGAGCAAG 1329  
Db 1412 TCAATTGACTCTCAGCACTCAATGGCGCTTTGAGCCTGTGAACATCAAGGCTGCCAGAC 1471  
QY 1330 CAGCAACCCCGATCTGTCTACCATCCAGCGGAGCTCAGCACCCACCACTACACAGG 1389  
Db 1472 AAGATCTCTGGCTGTCTCCCGCTGACC-----TCACCTCAGCTGCAGCCATGTACAGG 1525  
QY 1390 GCAGTCTCTGTCCCGGAGGATG-----GGCCAGCGCCCAAGTTTCCAGCTCACCAGTGG 1444  
Db 1526 GACCTGTCTATGCTCTGTGATGATGTCTCAGACAAATCCCAATGACCAACTCTCCAAATC 1585  
QY 1445 GCACCTGTCTAGCCCGCTGGTGGGCGGCGGACACATGTACACACAGCTCTCCACCTC 1504  
Db 1586 TGGACCCACTACCCAACTTGAATCAAAATGTAACACAGCTCAGGTGTCTGCTCACTCTC 1645  
QY 1505 -----TGAGGCGAGAGTGTCTCTCCCGCTCTCCACCC----- 1539  
Db 1646 AGGATGACCTTCCGAGTCTCATCCAACTGTCTCCCGAGATGACCCAGTCCCTGTGTAG 1705  
QY 1540 -----AGAACTACTTCCGCTCTCCCGCTCTCCCGGAGCACCAAGCA 1578  
Db 1706 AGAATGAGCCCTTAACCTGAAGAACGAGGCTTCGCAAGACAGACTGACCCATCTGTCA 1765  
QY 1579 TGACCTATGGGACCTTCACTCTCTGGGGCGGCTGATGATCTCTTAATACAGGTATCA 1638  
Db 1766 CAGCAATTTGGTACCTTCAACTCTCTTGGGGGTCACTCATCATCTTAATCAGGAGTAA 1825  
QY 1639 GCCTCTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCACGC 1698

Db 1826 GTTGTGATTCGCGTGGGCCATTTCTTCAGGGGAGAGTCTATGAAATGTATGTGACTG 1885  
QY 1699 TGCACAAGCCGGAAGAGCTGAGGTTGCCCTAGCTGGCTGTGAGACCTCTGCTGAGTCCCA 1758  
Db 1886 TACACAGGAAGAAATATGAGCGCCCATATGAGAGACTCTCAGACCTTACCTTACCCCTG 1945  
QY 1759 TCGTTAGCTGTGAGACCCCTGCGCTCTCTCACCGCGCAGTCACTCTGGCTTATGAGAC 1818  
Db 1946 TGTGTAGCTGTGGGCTCTCTGGAGCTTGTGTGACCGCGCTGTCTATCTCTCACTCTGATC 2005  
QY 1819 ACTGTGGGAGGCGCAGCCCTGACAGCTGGAGCCCTGCGCTCAAAAAGAGCTGTGTGGAG 1878  
Db 2006 ACTGTGCAGACCCACGAGGACTTGAAGATCCAGTCAAAAACGAGGAGTGTGGAG 2065  
QY 1879 GCAGCTGGAGGAGAGTGTGTGCTGCTGAGCGGAGGAGCGCCCTCCCACTCTACTACT 1938  
Db 2066 GACATGGGAGGATGTGTGGTG---GTTGGGGAGGAGAACTTCAAAACCCCTGTGTACA 2122  
QY 1939 GCCAGCTGGAGGCGAGTGTGCTGTCTTACCAGAGAGCTGGGCGCTTTGGCCCTGG 1998  
Db 2123 TTGAGTGTGATGAGAGGCTTGCATATCTTCACAGAGAACCTCAGTACCTATGCGCTGG 2182  
QY 1999 TGGGAGAGGCGCTCAGCGTGGCTGCGCGCAAGCGCTCAAGCTGTCTTCTTTGGCGCGG 2058  
Db 2183 TTGGGAGTCCACCAACAGAGAGCTGCCAAGCGTCTTAACTGGCCATCTTTGGGCGCC 2242  
QY 2059 TGGCTGTCACTCTCTCGAGTAAACATCGGGCTGTACTGCTGTGATGACACCCAGCATG 2118  
Db 2243 TCTGTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCTGTGATGACACAGAGT 2302  
QY 2119 CACTCAAGAGGCTGTGACGTGAGAAAGCAGCTGGGGGAGACAGCTGATCAGAGAGCAC 2178  
Db 2303 CCTGAAGAGAGTCTTACAACTGGAGAGCAATGGAGGAGAGCTCCTAGAGAACCCA 2362  
QY 2179 GGGTCTGTGCACTTAAAGGACAGTTACAAACCTGCGCTTATCCATCCAGATGTGCCA 2238  
Db 2363 AGGCTCTCTGTTTAAAGGAGAGTCCCAACCTGCGCTGTCTTATTCATGACATCGCC 2422  
QY 2239 GCTCCCTGTGGAGAGTAAAGCTCTTCTGACGTACAGAGAGATCCCTTTTATCACAAT 2298  
Db 2423 ATTCCCTCTGGAGAGCAATTTGCTGCTAAGTATCAGAAATTCATTTTACACATCT 2482  
QY 2299 GGAATGCGACGAGCGGTACTTGCATCTGCACCTTCAACCCTGGAGCGTGTACGCCCCAGCA 2359  
Db 2483 GGAGTGTCTTCAAGAAACCTTCACTGCACTTCACTCTGAAAAGACTCAGCTTAAACA 2542  
QY 2359 CTAGTGAACCTGCGCTGCAAGCTGTGGGTGTGGAGTGGAGGCGAGCGGAGAGCTTCA 2418  
Db 2543 CAGTGGAACTGGTTTGAACACTCTGTGTGCGGAGGTTGAAGGAGAGGCGAGATCTTCC 2602  
QY 2419 GCATCAACTTCAACATCAACAAGGACACAAAGTTTGTGAGCTGTGCTGTGAGAGT 2478  
Db 2603 AGCTCAACTGTACTGTGTGAGAGAACCTACTGGCATGCACTTACTCTCTGAGACCTG 2662  
QY 2479 AAGCGGGGTCCAGCGCTGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCTCATTTGGC 2538  
Db 2663 CTAGTACCATCACTGTCTACCGGACCAAGTGTCTTTCAGCATTCCTCTCGCTATCCGCG 2722  
QY 2539 AGAAGATAAATTCAGCGCTGAGACCCACCTGTAGCGGGGTGCCAGTGGCGGAGTCTGG 2598  
Db 2723 AGAAGCTATGAGAGCGCTGGATGCCCTCAACAGAGGCGCATGACTGGAGGATGTGG 2782  
QY 2599 CCCGAAACTCCACTGGAGAGCATCTCAGTCTTCTTGGCTTCAAGCGCCAGCCCCACAG 2658  
Db 2783 CCCATAACTCAACTGGAGAGTACTTGAATTTTGGCCACCAATTCGAGCCCCAATG 2842  
QY 2659 CCATGATCTCAACCTGTGGAGGCGCGCTCTCCCAAGCGCAACCTCAGCCAGCTGG 2718  
Db 2843 GCGTAATCTTGATCTTTTGGGAAGCACAGAACTTCCAGATGGAAACCTGAGCATGCTGG 2902  
QY 2719 CTGACAGAGTGGCTGCACTGGGCCAGCAGCGTGGCCCTCTTTCACAGTGTGGAGGCTG 2778

2903 CAGCGCTCTGGAGAAATGGGAGACATGACAGACAGTGGTGTACTTGGCAGCAGAGGAC 2962

Qy 2779 AGTSGCTGA 2786  
 ||| |||

Db 2963 AGTATTGA 2970

RESULT 4  
 LOCUS AK018177 3866 bp mRNA linear HTC 20-SEP-2003  
 DEFINITION Mus musculus adult male medulla oblongata cDNA, RIKEN full-length enriched library, clone:630415B02 product:TRANSMEMBRANE RECEPTOR UNC5H2 homolog [Rattus norvegicus], full insert sequence.

ACCESSION AK018177  
 VERSION AK018177.1 GI:12857775  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

REFERENCE  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

REFERENCE  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, E., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

REFERENCE  
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)

REFERENCE  
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)

REFERENCE  
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, K., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Db 600 AACAGCCAGTGGAACTGCAGCTGCAGAGCCTTCCCTGCCACGAGATCTACTTCAAGTGT 659  
 QY 318 AACGGGAGTGGGTGGCCAGGTGGACCAAGTGTATCGAGCGCAGACAGACGGAGCAGT 377  
 Db 660 AATGGCGAGTGGGTGAGCGAGATGACCAAGTGCACACAGAGAGCCTGGATGAGGCCACA 719  
 QY 378 GGTGAGCCGACATGATGAGGTGCCATATATGTCTCAAGGACGAGTGCAGAGGTGTTTC 437  
 Db 720 GGCTTGGGCTGCGAGAGTGCAGATCGAGGTGTCAAGGAGCAAGTGGAGGAACCTCTTC 779  
 QY 438 GGCTCGAGGAATACTGGTCCAGTGCCTGGATGAGTCTCTCGGCACCAACAAGAT 497  
 Db 780 GGCTCGAGGACTACTGGTCCAGTGGTGGCTTGAGCTTCTGGGAATCTACAAGAT 839  
 QY 498 CAGAAGCCTTACATCCGATPAGCCAGATTGCGAAGAACTTCGAGCAGAGGCGCTGGCC 557  
 Db 840 CGCGAGCCTACATCCGATTCCTACTTCCGCAAGAACTTTGACAGGAGCCTCTGGCC 899  
 QY 558 AAGGAGTGTCCCTGGAGCAGGACATGCTGCTGCCCTGCCGTCCACGGAGGCGATCCCT 617  
 Db 900 AAGGAGTACCTTGGATCATGAGTCTTCTGAGTGGCGGCCACCGAGGAGTGCCT 959  
 QY 618 CCAGCCAGTGGAGTGGCTCCGGAACGAGACCTGGTGGACCCGCTCCCTGGACCCCAAT 677  
 Db 960 GTGGTGAATGGATGGCTCAAGATGAAGATGCTCATTTGACCCCGCTCAGGACACTAAC 1019  
 QY 678 GTATACATACGCGGGAGCAAGCCTGGTGGTGGACAGGCGCGCCTTGTGTGACAGGCC 737  
 Db 1020 TTCTGTCTACCATTTGACCAACACTCATATCCGCGAGGCGGCTCTCAGACACGCCC 1079  
 QY 738 AACTACCTGCTGGCGCAGAACTGCTGGCAGTGCAGCGGCGCTCCGCTGCTGTC 797  
 Db 1080 AACTACCTGCTGGCGCAGAACTATCGTGCCCAAGCGCGGAGCAACGCGGCCACAGTC 1139  
 QY 798 ATCGCTACGTGAACGCTGGTGGTGGTGGACCGAGTGGTCCGCTGTGAGCGCCAGC 857  
 Db 1140 ATCGTATATGTAATGAGGCTGTCCAGCTGGGACAGTGGTCAACCTGTTCATTCGC 1199  
 QY 858 TGTGGGCGCGCTGGCGAAGAACGAGCGGAGTGCACCAACCGCGGCGCTCTCAACGGG 917  
 Db 1200 TGTGGCGAGGCTGGCAGAGCCTACTCGGACCTGCACCAATCCAGGCCCACTCAATGGA 1259  
 QY 918 GCGCTTCTGTGAGGCGCAGATGCTCATGACCGCACCGTCTCTCTCTGCTGTCTCT 977  
 Db 1260 GCGCGCTTCTGTGAGGAGCAGGCTTCCAGAGACAGTTCGACACCGTGTG---CCCA 1316  
 QY 978 GTGACCGGACCTGGAGCCGCTGGAGCAAGTGGTGGCTGGGCTGGAGTGCACCCAC 1037  
 Db 1317 GTGATGGAGCGTGGACCGAGTGGAGCAAGTGGTCTGCTGACGACAGAGTGTGCGCAC 1376  
 QY 1038 TGGCGGAGCGTGTGCTGTGACCCAGCACCCCGCAACGAGGAGGAGGAGTGCAGGCGC 1097  
 Db 1377 TGGCGGAGCGCGAGTGCATGCAACCGCACCCCGCAACGAGGAGGAGGAGTGCAGGCGG 1436  
 QY 1098 ACTGACCTGGACACCCGCACTGTACAGTGCACCTCTGTGACAGTGTCTTGCGCC-- 1155  
 Db 1437 ACCTACTTGACTCAAGAACTGCACTGATGAGTGTGCTGCTGATCAAGAACTCTA 1496  
 QY 1156 -----CTGAGGAGCGTGGCCCTCTATGTGGGCTC 1184  
 Db 1497 AACGACCTAAAGCCACCCCTGGAGACATCGGAGATGTGGCACTGTACGAGGCTT 1556  
 QY 1185 ATGCGCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1241  
 Db 1557 GTGTGGCGCTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1616  
 QY 1242 CGGAAGAGGAGGAGGAGTGCAGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298  
 Db 1617 AGAACTGCGGGACTTTCAGACAGGACATCACCGACTCTCTGCGGCGCTCTACTGCTGCT 1676  
 QY 1299 TTCCAGCCGCTGAGCATCAAGCCCGAGCAAGAGCAACACCCCACTGCT-----CACC 1352  
 Db 1677 TTCCACCTGTCAACTTCAAGACTGCAAGGCCCAACACCCCGAGCTCTGACCCCGTCC 1736

QY 1353 ATCCAGCCGAGCTCAGCACCACCACTACAGGCGAGTCTCTGTCCCGGCGAGGAT 1412  
 Db 1737 GCGCTTCCAGACCTTAAGGCGAGTGTGCACTTACCGGGGCGCTGTGTATGCGCTGCAG 1796  
 QY 1413 GGGCCCGAGCCCAAGTTCAGAGTCAACCAATGGGCACTGCT----- 1453  
 Db 1797 GACTCGCGCAGCAAGATCCCATGACTAATTCGCCCCCTGTGGATCCCTGCCAGCCTC 1856  
 QY 1454 -----C 1454  
 Db 1857 AAGATCAAGGTCTATAACTCCAGCAACATCGGTTCGTGGTCTGGCCTGGCTGATGAGCC 1916  
 QY 1455 AGCCCCCTGGGTGGCGCGCCACACACTGCACACAGTCTCTCCACCTCTGAGGCGGAG 1514  
 Db 1917 GACTCTGTGGTGTCTCTCCCGCGGCACTACCCAGGCAATTTCTCCCGGACACCCAT 1976  
 QY 1515 GAGTGTGTCTCCCGCCTCTCCACCCAG-----AATCTTCGCTCCCTGCGCCGAGGC 1568  
 Db 1977 TTCCTGCACCTGGCAGTGCAGCCTTGGTTGCCAGCACCTCTCGGCGCTTACCTCGGAC 2036  
 QY 1569 ACAGCAACATGACCTATGGGACCTTCACTTCTCGGGGCGGCTGATGATGATCCCTAAT 1628  
 Db 2037 CCAGCAGCAGTGTCAAGCGGCACTTTGGTGTGCTGGAGAAAGCTGAGCCTCCCGGC 2096  
 QY 1629 ACAGGTATCAGCCTCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGATC 1688  
 Db 2097 ACAGGGTCACTGTGTGGTACCAATGAGCCATTTCCCGAGGCAAGTTCTATGACCTG 2156  
 QY 1689 TACTCTACGCTGACAAAGCGGAGAGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1748  
 Db 2157 TATCTACATATCAACAAGGCGGAAAGCACCTCCCTCCACTTTTCAAGAGGTTCCGACAGTA 2216  
 QY 1749 CTGATGCCATGCTAGCTGTGAGCCCCCTGGGCTGCTGCTCACCCGCGAGTGCATCCTG 1808  
 Db 2217 TTGAGCCCCCTCGGTGACCTGTGGGCGCACAGGCGCTACTCTGTGCGGCGCTGCTGCTC 2276  
 QY 1809 GCTATGAGCACTGTGGGAGCGCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAG 1868  
 Db 2277 ACCGTGCCCACTGTGCTGAAGTCACTGCTGGAGACTGGATCTTTCACTCAAGACCCAG 2336  
 QY 1869 TCGTGGAGGCGAGCTGGGAGCAGATGTGCTGCACTGGGCGAGGAGGCGCCTCCAC 1928  
 Db 2337 GCCATCGGCGCACTGGG---AGGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2393  
 QY 1929 CTCTACTACTGCGAGCTGGAGGCGCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1988  
 Db 2394 CCCTGCTACTGCGAGCTGGAGGCTAAGTCTGCGCACATCTCTGCTGGACCACTGGGTTCC 2453  
 QY 1989 TTTGCCCTGTGGGAGGCGCCTCAGCGTGGCTGCGCGCAAGCGCCTCAAGCTGCTGTG 2048  
 Db 2454 TACGTATTCATGGCGAGTCTTACTCTGCTCTGCACTGCAAGCGGCTCCAGCTGGCCATC 2513  
 QY 2049 TTTGGCGCGGTGGCTGCACTCCCTCGAGTACAACTCCGGGTCTACTGCTGCTGCTGCTGCTGCT 2108  
 Db 2514 TTTGCCCGCAGCCTCTGCACTGCGCTGGAGTATAGCCTCAGGCTCACTGTCTGGAGGAC 2573  
 QY 2109 ACCACAGTGCACCTCAAGGAGTGTGCTGAGTGGAGAGCAGCTGGGCGGAGCAGCTGATC 2168  
 Db 2574 ACCTGTAGCACTGAAGGAGTCTTGGAGCTGGAGGAGTCTGGGTGGTCTGCTGCTGCTGCTGCT 2633  
 QY 2169 CAGGAGCCACCGGTCTGCACTTCAAGGACAGTTTACCAACCTGCGCCTATCCATCCAC 2228  
 Db 2634 GAGAGCCACAGCCTTGTCTTTAAGGACAGTTTACCAACCTAGGCTCTCCCTCCAT 2693  
 QY 2229 GATGCGCCAGCTCCCTGTGGAAGTGAAGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2288  
 Db 2694 GACATCCCGGCTGCGGCTGAGGAGCAAACTACTTGGCCAAAGTACCAAGAGATCCCTTC 2753  
 QY 2289 TATCAGTCTGAAATGGCAACGAGCGGTACTTGTGACTGCACTTCCCTTGGAGGCTGCT 2348  
 Db 2754 TACCACGTCTGAAATGGCGCAGAGCCCTGCACTTTCACCTTGGAGGAGCAT 2813

QY 2349 AGCCCCAGCACTAGTGAACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGGACGGG 2408  
 Db |||||  
 2814 AGCCTGGCTCCACGGAGTTACCTGTAAAGTCTGCTGGCGGAGGTGGAAGGGAAGGC 2873  
 QY 2409 CAGAGCTTCAGCATCACTTCAACATCAC--AAGGACACAAGTTTGTCTGAGCTGCTG 2465  
 Db |||||  
 2874 CAGATTTTCAGCTGTCACACAACTGTGGCCGAGACGCTGCTGCTCCCTGGATGCTCTC 2933  
 QY 2466 GCTCTGAGAGTGAAGCGGGGTCCAGCCTGTGGTGGSCCCAGTGCCTTCAAGATCCCC 2525  
 Db |||||  
 2934 TGCTCTGCCCGGCAATGCCATCACACCAGCTGGGACCTTATGCCCTTCAAGATACCC 2993  
 QY 2526 TTCCTATTCGGCAGAGATAAATTTCCAGCTTGAACCCACCTGTAGCGGGGTGCCGAC 2585  
 Db |||||  
 2994 CTGTCCATCGCCAAAGATCTGCAGCAGCTGTGACGCCGCCGACTCCCGGGCAACGAC 3053  
 QY 2586 TGGGGGACTCTGGCCCAAGAACTCCACTGACAGACCATCTACACTTCTTTGGCTTCCAAG 2645  
 Db |||||  
 3054 TGGAGGCTGTGGCGCAGAGCTGTCCATGACCGCTGACCTAAACTACTTTCGCCACCAA 3113  
 QY 2646 CCCAGCCCCACAGCCATGCTCAACTGTGGGAGCGGGGCACTTCCCCAAGCGCAAC 2705  
 Db |||||  
 3114 GCTAGTCCACAGGTGTCACTTAGACCTCTGGGAGCTCGGCAACAGGATGACGGGAC 3173  
 QY 2706 CTAGCCAGCTGCTGCAGCAGTGGCTGGACTGGGCCAGCAGACGCTGGCTCTTCA 2765  
 Db |||||  
 3174 CTCACAGCCTGGCCAGTGGCTTGGAGGAGATGGCCAAAGTGAGATGCTGTAGCCATG 3233  
 QY 2766 GTGTGGGAGCTGAGTGTCTGAG 2787  
 Db |||||  
 3234 GCCACAGATGGGATGTCTGAG 3255

RESULT 5  
 AY411747  
 LOCUS Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, 2532 bp DNA linear GSS 12-DEC-2003  
 DEFINITION Genomic survey sequence.  
 ACCESSION AY411747  
 VERSION AY411747.1 GI:39767715  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2532)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE 2 (bases 1 to 2532)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
 FEATURES  
 Location/Qualifiers  
 source 1..2532  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>2532  
 /locus\_tag="HCM4327"

Query Match 26.5%; Score 763.2; DB 29; Length 2532;  
 Best Local Similarity 60.0%; Pred. No. 7.6e-118;  
 Matches 1515; Conservative 0; Mismatches 873; Indels 138; Gaps 9;  
 QY 393 GAGGTCCGCAATTAATGTCTCAAGGCAGCAGGTTCGAGAGGTGTTGCGGCTGGAGGAATAC 452  
 Db |||||  
 13 GAGGTGAGATCGAGGTGTCCGCGCAGCAGGTGGAGAGCTCTTTGGGCTGGAGATTAC 72  
 QY 453 TGGTGCAGTGTGGCATGGAGCTCCTCGGGACACCAAGAGTCAAGAGGCTTACATC 512  
 Db |||||  
 73 TGGTGCAGTGTGGCATGGAGCTCCTCGGGACACCAAGAGTCAAGAGGCTTACATC 132  
 QY 513 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGAGCGCTGGCCAAAGAGGTTCCTG 572  
 Db |||||  
 133 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGAGCGCTGGCCAAAGAGGTTCCTG 192  
 QY 573 GAGCAGGGCATCTGTGCTGCCCTGCCGTCACCGAGGGCATCCCTCCAGCGGAGGTGGAG 632  
 Db |||||  
 193 GAGCAGGGCATCTGTGCTGCCCTGCCGTCACCGAGGGCATCCCTCCAGCGGAGGTGGAG 252  
 QY 633 TGGTCCCGAACGAGGACCTGTGGTGGACCCCTGCTGGACCCCAATGTATATACATCACGCG 692  
 Db |||||  
 253 TGGTCCCGAACGAGGACCTGTGGTGGACCCCTGCTGGACCCCAATGTATATACATCACGCG 312  
 QY 693 GAGCAGAGCTGTGGTGGCAGAGCGGCTGTGCTGACCGGCCAACTACACCTGCGTG 752  
 Db |||||  
 313 GAGCAGAGCTGTGGTGGCAGAGCGGCTGTGCTGACCGGCCAACTACACCTGCGTG 372  
 QY 753 GCCAAGAACTCTGTGGCAGCGCTGCCGCGCAGCGCTCGCTGCTGTCATCGTCTAGTGAAC 812  
 Db |||||  
 373 GCCAAGAACTCTGTGGCAGCGCTGCCGCGCAGCGCTCGCTGCTGTCATCGTCTAGTGAAC 432  
 QY 813 GGTGGTGTGTCAGCTGGACCGAGTGGTCCGCTGTGTCAGCGCCAGCTGTGGCGCGGTGG 872  
 Db |||||  
 433 GGTGGTGTGTCAGCTGGACCGAGTGGTCCGCTGTGTCAGCGCCAGCTGTGGCGCGGTGG 929  
 QY 873 CAGAAACGAGCGGAGCTGCACCAACCGCGGCTGTGTCAGCGCCAGCTGTGGCGCGGTGG 932  
 Db |||||  
 493 CAGAAACGAGCGGAGCTGCACCAACCGCGGCTGTGTCAGCGCCAGCTGTGGCGCGGTGG 552  
 QY 933 GGCAGAACTGTCCATGACCGCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 992  
 Db |||||  
 553 GGCAGAACTGTCCATGACCGCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 609  
 QY 993 AGCCCTGTGAGCAAGTGTGGTGGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1052  
 Db |||||  
 610 AGCGAGTGGAGCAAGTGTGTCAGCTGACACTGAGTGTGCTGCTGCTGCTGCTGCTGCTG 669  
 QY 1053 TGTCTGTGACCCAGCAGCGGCGGAGGAGGAGTGCAGGCGGCTGCTGCTGCTGCTGCTGCT 1112  
 Db |||||  
 670 TGTCTGTGACCCAGCAGCGGCGGAGGAGGAGTGCAGGCGGCTGCTGCTGCTGCTGCTGCT 729  
 QY 1113 CGCACTGTACCACTGACCTCTGTGTACACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1151  
 Db |||||  
 730 AAGAACTGTACCACTGACCTCTGTGTACACAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 789  
 QY 1152 -----GGCCCTGTGAGCAAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1196  
 Db |||||  
 790 CACCTGTGTGAGCAAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 849  
 QY 1197 GTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1256  
 Db |||||  
 850 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909  
 QY 1257 CTGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1313  
 Db |||||  
 910 TTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 969  
 QY 1314 ATCAAGCCAGCAAGAGCAGCAACCCCATCTGTCT-----CACCATCCAGCGGACCTC 1367  
 Db |||||  
 970 TTTAAGCGGCAAGGCGCAGCAACCCCATCTGTCTACACCCCTCTGTGTGTGTGTGTGTGT 1029  
 QY 1368 AGCACCCACCACTACAGGAGTCTGTGTCTCCCGGAGGATGTGGCCCGGACCCCAAG 1427

[illegible]

Db	2107	GAGCTCACCTGCAAGATCTGCTGGCGCAAGTGGGAAGGGGAGCCAGATATTCACGCTG	2116
Qy	2424	AACTTCAACATACCAAG--GACACAAGGTTTGCTGAGCTGTGGCTCTGGAGAGTGAA	2480
Db	2167	CATACCACTCTGCAGAGACACTGTGGCTCCCTGGACACTCTCTGCTGTGCCCTGGC	2226
Qy	2481	CGCGGGGTCCAGCCCTGTGTGGGCCCGAGTGCCTTCAAAGATCCCTTCTCATTCGCGAG	2540
Db	2227	AGCACTGTCAACCAACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAG	2286
Qy	2541	AAGATAATTTCAGCTTGGACCAACCTCTAGGCGGGGTGCGCACTGGGGGACTCTGGCC	2600
Db	2287	AAGATATGCAACAGCTTAGATGCCCCCACTCAGCGGGGCAATGCTGGCGGATGTTAGCA	2346
Qy	2601	CAGAACTCCACCTGGACAGCACTCTCAGCTTCTTTTGCTTCCAAGCCCAAGCCACAGCC	2660
Db	2347	CAGAGCTCTCTATGACCGGTACCTCAATTACTTTGCCACCAAGAGCGCCCAACGGGT	2406
Qy	2661	ATGATCCTCAACCTGTGGGAGCGCGCACTTCCCCCAACGGCACTCAGCAGCTGGCT	2720
Db	2407	GTGATCTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCTGGCG	2466
Qy	2721	GCAGCAGTGCCTGGACTGGGCCAGCCAGACGCTGGCCCTCTTCACAGTGTGGAGGCTGAG	2780
Db	2467	AGTGCCCTTGGAGAGATGGGCAAGAGTGAGATGCTGTGGCTGTGGCCACGACGGGAC	2526
Qy	2781	TGCTGA	2786
Db	2527	TGCTGA	2532
RESULT 6			
LOCUS	AY406492	2802 bp DNA linear	GSS 15-DEC-2003
DEFINITION	Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY406492		
VERSION	AY406492.1	GI:39762466	
KEYWORDS	GSS.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science	302 (5652), 1960-1963	(2003)
PUBMED	14671302		
REFERENCE	2 (bases 1 to 2802)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003)	Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..2802		
	/organism="Pan troglodytes"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		
	<1..>2802		
gene	/gene="UNC5C"		
ORIGIN	/locus_tag="HCM2575"		
Query Match	26.5%	Score 763.2;	DB 29; Length 2802;

Best Local Similarity 52.9%; Pred. No. 8e-118; Matches 1456; Conservative 0; Mismatches 1218; Indels 78; Gaps 7;			
QY	98	GCCCGGCTGTGGCCAGCGCTCCCTGGGCAATAGTCTCGCCGCTGGCTCCGGGCTCGGG	157
Db	56	GCTCGTGTAATCGTCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCAAGA	125
QY	158	TGCCAGCAGAGTGCCACCGTGGCCCAACCCAGTGCCTGGTGCACACCCGACCTGCTTCC	217
Db	126	TGATGACTTTTTCATGAATCCCAAGAACTTTTCCTTCGGATCCACCTGAGCCCTCTGCC	185
QY	218	CACTTCTCTGTGGAGCCGAGAGTGTGTAATCGTCAAGAACCAAGCCAGTGTCTGTGT	277
Db	186	ACATTTCTTATTGAACCTGAGGAAGTTATATTGTGAAGAATAAGCCGTGTGAACCTGTA	245
QY	278	GTGCAAGGCC-----GTGCCGCCACGACGATCTTCTCAAGTGCACAGGGGAGTGGT	331
Db	246	CTGTAAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	305
QY	332	GCGCCAGGTGGACACGCTGATCGAGCGCAGCACAGACGGGAGCAGTGTGAGCCGACAT	391
Db	306	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	365
QY	392	GGAGTCCGCAATTAATGTCTCAAGCGCAGAGTGCAGAAAGTGTTCGGGCTGGAGGAATA	451
Db	366	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	425
QY	452	CTGTGCCAGTGGTGCATGGAGCTCTCGGGACCAACAGAGTCAGAAAGCCCTACAT	511
Db	426	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	485
QY	512	CCGATAGCCAGATTGCGCAAGAACTTCGAGCAGAGCGCGCTGGCCAGGAGTGTCCCT	571
Db	486	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	545
QY	572	GGAGCAGGCAATCGTGTGCCCTGCTCCACCGAGGGCATCCCTCCAGCCGAGGTGGA	631
Db	546	GGAACAGGAAGTCTTACTCCAGTGTGGCCACCTGAAGGATCCAGTGGCTGAGNNNN	605
QY	632	GTGGCTCCGGAAAGAGGACCTGTGGACCGTCCCTGGACCCCAATGTATACATCAGCG	691
Db	606	NNNNNGAATAATGAAGACATAATGATCCGTTGAAGATCGGAATTTTATATTACTAT	665
QY	692	GGAGCAGCGCTGGTGTGCGACAGGCGCCGCTTGTGACACGCGCAACTACACCTGCGT	751
Db	666	TGATCAACACTCATATAAGCAGCGCCGACTCTGTGATCTGCAATTTATACCTTGT	725
QY	752	GGCCAAGAACATCGTGGCAGCTGCGCGAGCGCTCGCTGTGTCTCATGTCTACGTGAA	811
Db	726	TGCCAAAAACATTGTGCCAAGAGAAAGTACAACCTGCCACTGTATAGTCTATGTCAA	785
QY	812	CGGTGGGTGGTGCAGCTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGCGCGGCTG	871
Db	786	CGGTGGCTGGTCCACTGCAGCGAGTGGTGTGTGTAAACAGCCGCTGTGGACAGGGTA	845
QY	872	GCAGAACCGGACCGGAGTGCACCAACCGCGGCTCTCAACGGCGGCGCTTCTGTGTA	931
Db	846	TCAGAAACGTACAAGACTTGTACCAACCGGACCACTCAATGGGGGTGCCTTCTGTGA	905
QY	932	GGGCGAAGATGTCATGACCGCACCGTCTCTCTGTGTGTCTGTGTGGAAGCGCAGCTG	991
Db	906	AGGGCAGAGTGTGCGAAA---AATAGCTGTACTACGTTATGCCAGTGGATGGCAGGTG	962
QY	992	GAGCCGTGGAGCAAGTGTGCGCCCTGTGGCTGGACTGCACCCACTGGCGGAGCCGTGA	1051
Db	963	GAGCCATGGAGCAAGTGTGTACTTGTGAACTGAGTGCACCCACTGGCGCAGGAGGA	1022
QY	1052	GTGCTGTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACTGACAC	1111
Db	1023	GTGCAAGCGCCAGCCCCCAAGAAATGAGGACAGGACTGCGATGGCTCGTTTGAATC	1082
QY	1112	CCGAACTGTACAGTGAATCTGTGTATACAGTGTCTGTGCCCTGTAGGAGCTGGCCCT	1171
Db			
1083	CAAGAACTGCACTGATGGGCTTTTGCAATGCAGACTGCTCCTGATTCAGATGATGTTGCTCT	1142	
QY	1172	CTATGTGGG---CCTCATCGCCGTGGCGCTGTGCTGCTGCTGCTGCTGCTGCTCTCAT	1228
Db	1143	CTATGTTGGGATTTGTATAGCAGTATGCTTTGCTGGGATCTCTGTAGTGTGGGCTT	1202
QY	1229	CCTGTTTATTTCCGGAAGAAGAGGGGCTGGACTCAGATGTGGGTGACTGCTCAATTCT	1288
Db	1203	GTTTGTGTATCGGAAGAATCATCTGACTTTGAGTTCAGATATATTGACTCTCTGGCACT	1262
QY	1289	CACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCGCAGCAAAAGACAGACACCCCATCTGCT	1348
Db	1263	CAATGGGGGCTTTTCAGCTGTGAACATCAAG-----GCAGCAAGACAGATCTGCT	1313
QY	1349	CACCATCCAGCCGACCTCTAGCACCAACCCACCTAGCAGGGCAGTCTCTGCCCGGCA	1408
Db	1314	GGCTGTATCCCCAGACCTCAGCTCAGCTGAGCCATGTACAGAGCACTGTCTATGCCCT	1373
QY	1409	GGATGGGCCAGCCCAAGTTCAGCTCAACAAATGGGCACTGTCTAGCCCCCTGGGTGG	1468
Db	1374	GCATGACGTCTCAGACAAAAATCCCAATGACCAACTCTCCAATCTGGATCCACTGCCAA	1433
QY	1469	CGGCCGCCACACACTGCACACACAGCT-----CTCCCACTCTGAGGCCGAGGA	1516
Db	1434	CCTGAAATCAAGTGTACAAACCTCAGGTGTGTCACCCCCAAGATGACCTCTCTGA	1493
QY	1517	GTTGCTCTCCGCTCTCCACCCAGAACTACTTCGCTCCTGCTGCCCGAGCACCAGCA	1576
Db	1494	GTTTACGTCAAGTGTCTCCCTCAGATGACCCAGTCAATGTTGGAGAAATGAAGCCCTCAG	1553
QY	1577	CATGA-----CTATGGGACCTT	1594
Db	1554	CCTGAAGAACAGAGACTTAGAAGGCAGACTGATCCATCTGTACCGCAITTTGGCAGCTT	1613
QY	1595	CAACTCTCCGCGGCGGCTGATGATCCCTAATAAGGTATCAGCTTATCAGCTCTCTATCCCC	1654
Db	1614	CNNNTGCTGGNNNNACCTTATGTTCCCAATCAGGAGTCAAGTGTGATTTCCCGC	1673
QY	1655	AGATGCCATACCCCGAGGGAAGACTATAGATCTACTCAGCTGCACAAAGCCGGAAGA	1714
Db	1674	TGGGGCAATCCCAAGGAGAGTCTACGAAATGTATGTACTGTACAGAGAAAGAAAC	1733
QY	1715	CGTGAAGTTCCTCCTAGCTGCTGCTCAGACCTGTGAGTCCCATCGTGTGCTGTGAGC	1774
Db	1734	TATGAGGCCACCCATCGATGACTCTCAGACATTTTGAACCTGTGTGAGTGTGGGC	1793
QY	1775	CCCTGCGCTCTGCTCACCGCGCAGTCACTCTGCTATGACCACTGTGGGAGGCCAG	1834
Db	1794	CCCAGAGCTCTGCTCACCGCGCCGCTGCTCTTATGCTATGCTATGCTGCTGAGTGTGAGC	1853
QY	1835	CCCTGACGTGAGCCTCGGCTCAAAAGCAGTCTGCGAGGGCAGTGTGGGAGCAGGA	1894
Db	1854	TACCGAGGACTGGAAATATCTGCTCAAGAACCCAGGACAGCAGGACAGTGGG--AGGA	1910
QY	1895	TGTGCTGCACCTGGCGAGGAGCGCCCTCCCACTCTACTGCTGCTGAGTGTGAGGCCAG	1954
Db	1911	TGTGTTGTTGGTTCGGGGAGGAAACTTCAACCCCTCTACTATTCAGTGTGATCAG	1970
QY	1955	TGCTGCTACGTCTTTCACCGAGCAGTGGCGCTTTCCTGCTGCTGGGAGAGCCCTCAG	2014
Db	1971	GGCTTGCACATCTCTCAGAGAACTCAGCACCCTGAGCCCTGGTAGGACATTTCCACCAC	2030
QY	2015	CGTGTGCTGCCCAAGCGCTCAAGCTGCTTCTGTTTTCGCGCGGTGGCTGCACTCCCT	2074
Db	2031	CAAGCGGTGCGAAGCGCTCAAGCTGGCCATCTTTGGGCCCTGTGCTGCTCTCTGCT	2090
QY	2075	CGAGTACAACTCCGGTCTACTGCTGATGACACCAACGATGACACTCAAGGAGTGT	2134
Db	2091	GGAGTACAGACTCCGAGTCTACTGTCTGATGACACCCAGGATGCCCTGAAGAAATTTT	2150
QY	2135	GCAGCTGAGAGAGCAGCTGGGGGACAGTATCCAGAGCCACCGGTCTCTGACTTCAA	2194
Db	2151	ACATCTTGAGAGACAGATGGGAGGACAGCTCCTAGAGAAACCTTAGGCTCTTATTATAA	2210



```

12195 GGACAGTATCCACAACTGGCGCTATCCATCCAGATGTGCCAGAGTCCCTGTGGAGAG 2254
12196 |||||
12197 Db AGCAGACCCACAACTGGCGCTGTCAATTACAGATATGCCCATTCCTCTGGAGAG 2270
12198 |||||
12199 QY TAAGCTCCCTGTGTCAGTACAGAGATCCCTTTTATCATCTGGAATGGCAGCGAG 2314
12200 |||||
12201 Db CAAATGCTGGCTAAATATCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2330
12202 |||||
12203 QY GTACTTGGCACTGCACCTTCACTCTGGAGCGTGTGAGCGGCGAGCTAGTACCTGGCGCTG 2374
12204 |||||
12205 Db NAACCTGCATGCACCTTCACTCTGGAAGATTAGCTGAACACAGTGCAGCTGGTTG 2390
12206 |||||
12207 QY CAAGCTGTGGGTGTGAGGTGAGGGCGAGCGGCGAGCTTCCAGCATCACTTCAACAT 2434
12208 |||||
12209 Db CAAACTCTGTGTGGCGCAGGTGGAAGGAGGAGGAGCTTCCAGCTCAACTGCACCGT 2450
12210 |||||
12211 QY CACCAAGGACACAGGTTTCTGAGCTGTGGCTCTGAGAGTGAAGCGGGGTCCCGAGC 2494
12212 |||||
12213 Db GTGAGAGGAACCTACTGGCATCGATTGCGCTGCTGGATCTCGGACACCATCACCAC 2510
12214 |||||
12215 QY CTTGGTGGGCGCCAGTCCCTTCAAGATCCCTTCTCATTTGGGAGAGATAATTTCCAG 2554
12216 |||||
12217 Db GGTACGCGGCGCCAGTCTTTTTCAGCATCCCTCTCCCTATCCGGCAGAGCTCTGTAGCAG 2570
12218 |||||
12219 QY CTTGGACCCACCTGTGTAGGCGGGGTGCGGACTGCGGACTCTGCGCCAGAACTCCACCT 2614
12220 |||||
12221 Db CTTGGATGCCCCCAGACGAGAGGCCATGACTGAGGATGCTGGCCCATAGCAGNNNNNN 2630
12222 |||||
12223 QY GGACAGCATCTCAGCTTTTGTCTCAGCCCGAGCGCCATGATGATCTCTCAACCT 2674
12224 |||||
12225 Db NNNNNGTACTTGAAATTTGTCACCAATCCAGCCCACTGGCGTAACTCTGGATCT 2690
12226 |||||
12227 QY GTGGAGCGCGGCACCTTCCCAACGCGCACTCAGCCAGCTGCTGCGAGCAGTGGCTGG 2734
12228 |||||
12229 Db TTGGGAAGCAGAACTTCCAGATGGAACCTCAGCATGTGCGCAGCTGCTTGGAGA 2750
12230 |||||
12231 QY ACTGGGCGAGCAGCGTGGCTCTTTCAGAGTGTGGAGCTGAGTGCATGA 2786
12232 |||||
12233 Db AATGGGAAGCATGAAACGGTGTCTCTTAGCAGCAGAGGGCAGTATTAA 2802
12234 |||||

RESULT 7
BI758231 1034 bp mRNA linear EST 25-SEP-2001
LOCUS 603029876F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200171 5',
DEFINITION mRNA sequence.
ACCESSION BI758231
VERSION BI758231.1 GI:15749809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1034)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgepb@remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1501 row: g column: 20
High quality sequence stop: 793.
Location/Qualifiers
1..1034
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source
```

ORIGIN		Query Match	26.2%;	Score 756.2;	DB 12;	Length 1034;
		Best Local Similarity	89.7%;	Pred. No. 7.5e-117;		
		Matches	905;	Conservative	0;	Mismatches 88; Indels 16; Gaps 8;
QY	191	GCCTGTGTCACACCGGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACAT	250			
Db	1	GCCTGTGTCACACCGGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACAT	60			
QY	251	CGTCAAGAAACAAGCAGTGTCTTGTGTCAGAGCGCGTCCCGCCACGAGATCTTCTT	310			
Db	61	CGTCAAGAAACAAGCAGTGTCTTGTGTCAGAGCGCGTCCCGCCACGAGATCTTCTT	120			
QY	311	CAAGTGCACGGGAGTGGGTGCGCAGGTGGACCATGATCGAGCGCAGCAGACGCG	370			
Db	121	CAAGTGCACGGGAGTGGGTGCGCAGGTGGACCATGATCGAGCGCAGCAGACGCG	180			
QY	371	GAGCAGTGTGTGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGGCGAGGATCGAGAA	430			
Db	181	GAGCAGTGTGTGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGGCGAGGATCGAGAA	240			
QY	431	GGTGTTCGGCTGAGAGAAATATCTGTGTCAGTGGTGGCATGGAGTCTCTCGGCGACAC	490			
Db	241	GGTGTTCGGCTGAGAGAAATATCTGTGTCAGTGGTGGCATGGAGTCTCTCGGCGACAC	300			
QY	491	CAAGAGTCAGAGCGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC	550			
Db	301	CAAGAGTCAGAGCGCTTACATCCGATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC	360			
QY	551	GCTGCCCAAGAGAGTGTCTCTGGAGCAGGGCATCGTGTGCTGCTGCCCTGCCCGAGGG	610			
Db	361	GCTGCCCAAGAGAGTGTCTCTGGAGCAGGGCATCGTGTGCTGCTGCCCTGCCCGAGGG	420			
QY	611	CATCCCTCCAGCCGAGGTGGAGTGGTCCGGAACAGGACCTGTGTGACCCGCTCCCTGGA	670			
Db	421	CATCCCTCCAGCCGAGGTGGAGTGGTCCGGAACAGGACCTGTGTGACCCGCTCCCTGGA	480			
QY	671	CCCCAATGTATACATCACGGGGAGCAGCCTGTGTGTGGGACAGGCGCCGCTTGTCTGA	730			
Db	481	CCCCAATGTATACATCACGGGGAGCAGCCTGTGTGTGGGACAGGCGCCGCTTGTCTGA	540			
QY	731	CACGCGCAACTACACCTGCGTGGCCAAAGAAATCATCGTGGCAGCTGCGCGCAGGCTCCGC	790			
Db	541	CACGCGCAACTACACCTGCGTGGCCAAAGAAATCATCGTGGCAGCTGCGCGCAGGCTCCGC	600			
QY	791	TGCTGTATCTCTACGTGAAACGGTGGTGTGTGA- CGTGGACCGAGTGGTCCCTCTGCA	849			
Db	601	TGCTGTATCTCTACGTGAAACGGTGGTGTGTGTGA- CGTGGACCGAGTGGTCCCTCTGCA	659			
QY	850	GCGCCAGCTGTGGCGGGCTGGCAGAAACCGAGCGGAGCTGCGACCAACCGCGCGCTC	909			
Db	660	GCGCCAGCTGTGGCGGGCTGGCAGAAACCGAGCGGAGCTGCGACCAACCGCGCTC	719			
QY	910	TCAACGGGGCGCTTTTGTGAGGGGAGAAATGTCCATGACCGCACCGTCTCTCTGCG	969			
Db	720	TCAACGGGGCGCTTTTGTGAGGGGAGAAATGTCCATGACCGCACCGTCTCTCTGCG	774			
QY	970	TTGTCTGTGTGACGGGAGCTGGACCGCTGGAGCAAGTGTGCGGCTGT--GGGCTGGA	1027			



Db	775	CTGTGCCAGTGAAGCGCAGCTGTATGACACGTGTAGAGCCAGTGTGTGGCTGTGTGGCTGTGGC	833
Qy	1028	CTGACCCCACT--GGCGAGCGCGTGAAGTCTGTACCCAGCAGCCCGCAACGGAGGGAG	1085
Db	835	TTGACCCCACTTGGCGGAGCGGAGTGTCTGTARACCCAGCAGCCCGCAGCGAGGGG	894
Qy	1086	GAGTG---CGAGGGCACTGACCTGTGACACCCCGCAACTGTACCAGTGACCTCTGTGTACAC	1142
Db	895	AGGTGTGCCAGGTCACTGGACCTGGCACCCCGGA-TGGTCCAGTGAGCTCTGTGT-CCC	952
Qy	1143	AGTGCTTCTGCGCCCTGAGGAGTGGCCCTCTATGTGGGCCCTCATGCGG	1191
Db	953	ACGGTTCCTGCGCCCTGAGGACTTGGCCCTCTATGTGGGCCCTCATCCCG	1001
RESULT 8			
AY4111749			
LOCUS	AY4111749		
DEFINITION	Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,	2532 bp	DNA linear GSS 12-DEC-2003
ACCESSION	AY4111749		
VERSION	AY4111749.1	GI:39767717	
KEYWORDS	GSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1. (bases 1 to 2532) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2. (bases 1 to 2532) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
AUTHORS	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
TITLE	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
JOURNAL	Location/Qualifiers		
COMMENT	1. .2532 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>2532 /locus_tag="HCM4327"		
FEATURES	source		
gene			
ORIGIN			
Query Match	25.2%;	Score 726.4;	DB 29; Length 2532;
Best Local Similarity	59.1%;	Pred. No. 1.2e-111;	
Matches 1492;	Conservative	0; Mismatches 896;	Indels 138; Gaps 9;
Qy	393	GAGGTCCGCAATTAAATGTCTCAAGCGACAGGTTCGAGAAAGTGTTCGGGTGGAGGAATAC	452
Db	13	GAGGTTCAGATCAGGTGTCAAGCGACAGTTCGAGAAAGTGTTCGGGTTCGAGGAATAC	72
Qy	453	TGCTGCAGTGCCTGSCATGGAGCTCTCTGGGCACACCAAGAGTCAGAAAGCCTACATC	512
Db	73	TGCTGCAGTGCCTGSCATGGAGCTCTCTGGGCACACCAAGAGTCAGAAAGCCTACATC	132
Qy	513	CGCATAGCCAGATTTCGCGAAGAACTTCGAGCAGGAGCCGCTGGCCNAGAGGTGTCCTG	572
Db	133	CGCATAGCCAGATTTCGCGAAGAACTTCGAGCAGGAGCCCTGGCCNAGAGGTGTCCTG	192
Qy	573	GAGCAGGCACTGTGTGTCGCCCTGCGCTCCACCGAGGGGATCCCTCCACCGAGGTGGAG	632

QY	1584	TATGGACCTTCAACTTCTCTGGGGCCCGCTGATGATCCCTTAATACAGGTATCAGCCTC	1643
Db	1330	AGCGGCACTTTGGTGGCTGGAGGAAAGCTGAGCCTCCCGGACAGAGGGGTGAGCCTG	1389
QY	1644	CTCATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCGTGCAC	1703
Db	1390	TTGGTACCAAAATGAGCCATTCCTCCAGGGCAAGTTCTATGACCTGTATCTACATATCAAC	1449
QY	1704	AAGCCGAAAGACGTGAGGTTGCCCTAGCTGGCTGTGACACCCCTGCTGAGTCCCATCGTT	1763
Db	1450	AAGCCGAAAGACCTCCACATTTTCAGAAAGTTCCACAGACAGTAITGAGCCCTCGGTG	1509
QY	1764	AGCTGTGGAACCCCTCGCGTCTGCTACACCGGCGAGTCACTCTGGCTATGAGACACTGT	1823
Db	1510	ACCTGTGGGCCCCACAGAGCCTACTCTGTGCGCCCTGTGCTCCTACCGTGCACCCTGT	1569
QY	1824	GGGAGCCAGCCCTGACAGCTGGAGCTGGCTCAAAAGAGTCTGCGAGGGGAGC	1883
Db	1570	GCTGAAGTCATCGCTGGAGACTGGATCTTTTCAGCTCAAGACCCAGGCCCATTCAGGGCCAC	1629
QY	1884	TGGAGCAGATGTGTGCACTTGGGCGAGGAGCGCCCTCCACCTCTACTACTGCCAG	1943
Db	1630	TGGG---AGAGGTGGTGACTTGGATGAGAGACCCTCAACACACCCTCTACTGCCAG	1686
QY	1944	CTGAGGCCAGTGCCTGCTACTGCTTCCAGAGAGCTGGGCGCTTGTGCGCTGGTGGGA	2003
Db	1687	CTGAGGCTAAGTCTCTGCCACATCTCTGTCGACAGCTGGGTACCTACGTATTATGAGGC	1746
QY	2004	GAGGCTCAGCGTGGTGGCGGCAAGCGCTCAAGCTGCTTGTGTCGCGGTGGCC	2063
Db	1747	GAGTCTACTCTGCTGTCAGTCAAGCGCTCAGCTGGCTGCGCACTTCGCCACGCCCTC	1806
QY	2064	TGCACCTCCCTCGAGTACAAATCCGGGTCTACTGCTGATGACACCCAGATGCACTC	2123
Db	1807	TGCACCTCCCTGGAGTATAGCTCAGGCTCTACTGCTGGAGGACACACCTGTAGCACTG	1866
QY	2124	AAGAGTGTGTGAGCTGGAGAGACAGCTGGGGGAGAGCTGATCCAGGAGCCAGGGTC	2183
Db	1867	AAGAGTGTGTGAGCTGGAGAGGACTCTGGTGGCTACTTGGTGGAGGCCCAAGCCT	1926
QY	2184	CTGCACCTCAGGACAGTACCAACCTGGCTATCCATCCAGATGTGCCAGTCC	2243
Db	1927	TTGCTCTTTAAGAGAGTTTACCAACCTAGCCTCTCTCTCCATGATATCCCCCATGCC	1986
QY	2244	CTGTGGAAGTATAGCTCTTGTGACCTACAGGAGATCCCTTTTATCATATGGAAT	2303
Db	1987	CAGTGGAGGACAACTACTGGCCAAAGTACCAGGAGATTCCTTCTACCACTCTGGAAT	2046
QY	2304	GGCAGCAGCGGTACTTGCACTGACACCTTACCTTGGAGCGTGTGAGCCCGACACTAGT	2363
Db	2047	GGCAGCCAGAGCCCTGCACTGCACTTTCACCTGGAGAGGATAGCCTGGCCTCCACG	2106
QY	2364	GACCTGGCTCAGAGCTGTGGGTGTGGCAGTGGAGGGCGAGCGGACAGCTTCAGCATC	2423
Db	2107	GAGTTCACTTATAGCTTCTGCTGGGAGAGTGCAGAGGGGAGGCCAGATTTTCCAGCTG	2166
QY	2424	AACCTTCAACATCAC---AAGGACACAGGTTTGTGCTGAGCTGTGCTGTCTGCTGCTGCTG	2480
Db	2167	CACACAAAGTTGGCCGAGACGCTGTGCTGCTCCCTGGATGCTCTCTGCTGCTGCTGCTG	2226
QY	2481	CGGGGGTCCAGCGCTGGTGGGCCCCAGTGGCTTCAGATCCCTTCTCATTCGGCAG	2540
Db	2227	AATGCCATCACACCCAGCTGGGACCTTATGCTTTCAAGATACCCCTGTCTCATTCGCCAA	2286
QY	2541	AAGATAATTTCCAGCTGGAACCACTGTAGGGGGGTGGCGAATGGCGACTCTGGCC	2600
Db	2287	AAGATCTGCAGAGCTGGAAGCCCCCAACTCCCGGGGCAAGCACTGGAGGCTGTGGCG	2346
QY	2601	CAGAACTCCACTGGAGAGCACTCTAGCTTTCTTTTGGCTTCCAAAGCCAGCCCAAGCC	2660
Db	2347	CAGAAGCTGTCTCATGACCGGTACTCTAACTACTCTCGCCACCAAAAGCTAGTCCCAAGT	2406
QY	2661	ATGATCTCAACTGTGGAGCGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGCT	2720
Db	2407	GTCATCTTAGACCTCTGGGAGCTCGGCAACAGGATGACGGGACCTCAACAGCTGGCC	2466
QY	2721	GCAGCAGTGGCTGAGCTGGGCGCAGCAGCGTGGCTCTTTCACAGTGTGGAGGCTGAG	2780
Db	2467	AGTGCCTTGGAGGAGATGGCAAGAGTGTGATCTGTAGCCATGGCCACAGATGGCGAT	2526
QY	2781	TGCTGA	2786
Db	2527	TGCTGA	2532
RESULT 9			
BX364574			
LOCUS			
DEFINITION			
BX364574 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens			
cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.			
ACCESSION			
BX364574			
VERSION			
BX364574.1 GI:30368812			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 909)			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
COMMENT			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 3529.f For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS1AC006ZF06QP1&cluster=3529.f. Contact :			
Peng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS1AC006ZF06QP1.			
FEATURES			
Location/Qualifiers			
1..909			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DC023YB10"			
/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"			
/clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo (dr)			
primer. Five prime end enriched, double-strand cDNA was			
digested with Not I and cloned into the Not I and EcoR V			
sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 739; Conservative 4; Mismatches 7; Indels 1; Gaps 1;			
QY	1	AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCGCTCCCGCCCGCGGGGCGCGCGCG	60
Db	160	AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCGCTCCCGCCCGCGGGGCGCGCGCG	219
QY	61	GCCCGCGCGCTCCCGCGCGCGCGCGCGCTGGCGCGCTGGCGCGCGCGCGCGCTCC	120
Db	220	GCCCGCGCGCTCCCGCGCGCGCGCGCGCTGGCGCGCTGGCGCGCGCGCGCGCTCC	279
QY	121	TGGGCATAGTCTCCCGCGCTGGCTCCCGCGCTGGGTGCCAGAGTGCACCGTGG	180
Db	280	TGGGCATAGTCTCCCGCGCTGGCTCCCGCGCTGGGTGCCAGAGTGCACCGTGG	339
QY	181	CCACCCAGTGTGGTGCACACCGGACCTGCTTCCCACTTCTGTGTGGAGCCCGAGG	240



RESULT 11	BC033727	1532 bp	mRNA	linear	HTC 19-NOV-2003
LOCUS	IMAGE:5166762				HTC 19-NOV-2003
DEFINITION	Homo sapiens netrin receptor Unc5h1, mRNA (cDNA clone IMAGE:5166762), containing frame-shift errors.				
ACCESSION	BC033727	1	GI:21707230		
VERSION	HTC				
KEYWORDS	Homo sapiens (human)				
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
ORGANISM	1 (bases 1 to 1532)				
REFERENCE	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uudin, T.B., Toshikiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Hale, J.A., Gunatratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shavchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalek, U., Smallos, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
JOURNAL	2238257				
MEDLINE	12477932				
PUBMED	2 (bases 1 to 1532)				
REFERENCE	Strausberg, R.				
AUTHORS	Direct Submission				
TITLE	Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
JOURNAL	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
REMARK	Contact: MGC help desk				
COMMENT	Email: <a href="mailto:cgabs@mail.nih.gov">cgabs@mail.nih.gov</a>				
	Tissue Procurement: Life Technologies, Inc.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),				
	Gaithersburg, Maryland;				
	Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a>				
	Contact: <a href="mailto:nisc.mcg@hri.nih.gov">nisc.mcg@hri.nih.gov</a>				
	Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Iaric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripp, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.				
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>				
	Series: IRAK Plate: 68 Row: 1 Column: 2				
	This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction				
	This clone has the following problem: frame shifted.				
FEATURES	Location/Qualifiers				
source	1..1532				
	/organism="Homo sapiens"				

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5166762"
/tissue_type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

```

# ORIGIN

Query Match	24.7%	Score	711.2	DB 11	Length	1532			
Best Local Similarity	83.3%	Pred. No.	3.3e-109						
Matches	910	Conservative	0	Mismatches	8	Indels	174	Gaps	2
Qy	378	GGTGAGCCGACCATGGAGGTC	CCGCAATTAATGCTCTCAAGGCAGCAGGTCGAGAAAGGTGTTCC	437					
Db	612	GGGCTGCCACCATGGAGGTC	CCGCAATTAATGCTCTCAAGGCAGCAGGTCGAGAAAGGTGTTCC	671					
Qy	438	GGGCTGGAGGAATACTGGTGC	AGTGCCTGGCATGAGGTCCTCGGGCACCAAGAGT	497					
Db	672	GGGCTGGAGGAATACTGGTGC	AGTGCCTGGCATGAGGTCCTCGGGCACCAAGAGT	731					
Qy	498	CAGAAGCCTACATCCGCATAG	CCAGATTCGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC	557					
Db	732	CAGAAGCCTACATCCGCATAG	CCAGATTCGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC	791					
Qy	558	AAGGAGGTGCCCTGGAGCAGG	GCATCGTGCCTGCCCTCCACCGAGGGCATCCCT	617					
Db	792	AAGGAGGTGCCCTGGAGCAGG	GCATCGTGCCTGCCCTCCACCGAGGGCATCCCT	851					
Qy	618	CCAGCCGAGGTGGAGTGGCT	CCGGAACGAGGACCTGGTGGACCCGTCCTTGGACCCCAAT	677					
Db	852	CCAGCCGAGGTGGAGTGGCT	CCGGAACGAGGACCTGGTGGACCCGTCCTTGGACCCCAAT	911					
Qy	678	GTATACATCAGCGGGAGCAG	CGCTGGTGGTGGACAGGCCCGCTTGTGACACGGCC	737					
Db	912	GTATACATCAGCGGGAGCAG	CGCTGGTGGTGGACAGGCCCGCTTGTGACACGGCC	971					
Qy	738	AACATACACCTCGTGGCCAA	GAACATCGTGGCAGCTCGCGCAGCGCTCCGCTGCTGTC	797					
Db	972	AACATACACCTCGTGGCCAA	GAACATCGTGGCAGCTCGCGCAGCGCTCCGCTGCTGTC	1031					
Qy	798	ATCGTCTACGTGAACCGTGG	TGGTGGTGGACCGAGTGGTCCGCTCTGCAGCGCCAGC	857					
Db	1032	ATCGTCTAC		1040					
Qy	858	TGTGGCGCGGCTGGAGAAAC	GGAGCCGAGCTGCACCAACCGCGCGCTCTCAACGGG	917					
Db	1041			1040					
Qy	918	GGCGCTTCTGTGAGGGGCA	GAATGCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCT	977					
Db	1041			1040					
Qy	978	GTGACCGCAGCTGGAGCCCG	TGGAGCAAGTGGTCGGCTGTGGCTTGGACTGCACCCAC	1037					
Db	1041	GTGACCGCAGCTGGAGCCCG	TGGAGCAAGTGGTCGGCTGTGGCTTGGACTGCACCCAC	1100					
Qy	1038	TGGCGGAGCCGTGAGTCTCT	GTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCACGGGC	1097					
Db	1101	TGGCGGAGCCGTGAGTCTCT	GTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCACGGGC	1160					
Qy	1098	ACTGACCTGGACACCCGCACT	GTACAGTGACCTCTGTGTACACAGTGTCTTGGCCCT	1157					
Db	1161	ACTGACCTGGACACCCGCACT	GTACAGTGACCTCTGTGTACACAGTGTCTTGGCCCT	1220					
Qy	1158	GAGGACGTGGCCCTCTATGT	GGGCTCATGCCCGTGGCCGTCTGCCCTGGTCTCTGCTG	1217					
Db	1221	GAGGACGTGGCCCTCTATGT	GGGCTCATGCCCGTGGCCGTCTGCCCTGGTCTCTGCTG	1280					
Qy	1218	CTTGCTCTCATCTCGTTTAT	TGGCGGAGAGAGAGGGGCTGGACTCAGATGTGGCTGAC	1277					
Db	1281	CTTGCTCTCATCTCGTTTAT	TGGCGGAGAGAGAGGGGCTGGACTCAGATGTGGCTGAC	1340					

```

QY 1278 TCGTCCATCTCACCTCAGGCTCCAGCCGTGAGCATCAAGCCAGCAAGCAGACAAC 1337
Db 1341 TCGTCCATCTCACCTCAGGCTCCAGCCGTGAGCATCAAGCCAGCAAGCAGACAAC 1400
QY 1338 CCCCATCTGTCTACCATCCAGCCGAGCTCAG---CACCAACCAACCACTACCAAGGCGAGT 1394
Db 1401 CCCCATCTGTCTACCATCCAGCCGAGCTCAGCAACCAACCACTACCAAGGCGAGT 1460
QY 1395 CTCGTCTCCCGGAGGATGGCCCGAGCCCAAGTTCAGCTCCCAATGGCCACTCTCTC 1454
Db 1461 CTCGTCTCCCGGAGGATGGCCCGAGCCCAAGTTCAGCTCCCAATGGCCACTCTCTC 1520
QY 1455 AGCCCCCTGGGT 1466
Db 1521 AGCCCCCTGGGT 1532

RESULT 12
BI818609
LOCUS
DEFINITION 60303362P1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
mRNA sequence.
ACCESSION BI818609.1 GI:15929902
VERSION BI818609
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: scapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11434 row: 1 column: 16
High quality sequence stop: 744.
FEATURES
Location/Qualifiers
source 1..818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5174559"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/notes="Organ: Pooled brain, lung, testis; Vector:
pCMV-SpOR6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcorV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 24.6%; Score 710; DB 12; Length 818;
Best Local Similarity 97.1%; Pred. No. 4e-109;
Matches 744; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 84 GCCATGCCCTCGGCCCGCGCTGTGGCCAGCGCTCTCTGGGCATGCTCTCGCGCTTGG 143
Db 1 GCCATGCCCTCGGCCCGCGCTGTGGCCAGCGCTCTCTGGGCATGCTCTCGCGCTTGG 60
QY 144 CTCGCGGGCTCGGGTGGCCAGCAGAGTGCCACCGTGGCCCAACCCAGTGCTGTGTCACAAC 203

```

```

Db 61 CTCGCGGGCTCGGGTGGCCAGCAGAGTGCCACCGTGGCCAAACCCAGTGCTGTGTCACAAC 120
QY 204 CCGGACCTGTTTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAG 263
Db 121 CCGGACCTGTTTCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACAAG 180
QY 264 CCAAGTCTGTTTGTGTCAAGGGCGGTGCCCGCCACGAGATCTTCTTCAAGTGAACGGG 323
Db 181 CCAAGTCTGTTTGTGTCAAGGGCGGTGCCCGCCACGAGATCTTCTTCAAGTGAACGGG 240
QY 324 GAGTGGGTGGCCAGAGTGGACCACTGATCGAGCGCAGACAGACGGAGCAGTGTGTGAG 383
Db 241 GAGTGGGTGGCCAGAGTGGACCACTGATCGAGCGCAGACAGACGGAGCAGTGTGAG 300
QY 384 CCGACCATGAGGTCCGCATTAAATGTCTCAAGGCAGCAGGTTCGAGAAAGTGTTCGGGCTG 443
Db 301 CCCACCATGAGGTCCGCATTAAATGTCTCAAGGCAGCAGGTTCGAGAAAGTGTTCGGGCTG 360
QY 444 GAGGATACCTGGTGGCCAGTGGGCATGGAGCTCTCGGGCACCACCAAGAGTCAAGAG 503
Db 361 GAGGATACCTGGTGGCCAGTGGGCATGGAGCTCTCGGGCACCACCAAGAGTCAAGAG 420
QY 504 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGAG 563
Db 421 GCCTACATCCGCATAGCCATTATTTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGAG 480
QY 564 GTGTCCCTGGAGCAGGGGCATCGTGTGCTGCCCTGCCCTCCACCGAGGGGCATCCCTCCAGCC 623
Db 481 GTGTCCCTGGAGCAGGGGCATCGTGTGCTGCCCTGCCCTCCACCGAGGGGCATCCCTCCAGCC 540
QY 624 GAGTGGAGTGGCTCCGGAAACGAGGACCTGTGTGGACCCCTCCCTGGACCCCAATGTATAC 683
Db 541 GAGTGGAGTGGCTCCGGAAACGAGGACCTGTGTGGACCCCTCCCTGGACCCCAATGTATAC 600
QY 684 ATACGCGGGAGCAGACAGCCCTGTGTGCGACAGCCCGCTTGTGTGACACGGCCAACTAC 743
Db 601 ATACGCGGGAGCAGACAGCCCTGTGTGCGACAGCCCGCTTGTGTGACACGGCCAACTAC 660
QY 744 ACTGTGCTGCCCAAGAACATCGTGGCAAGTTC-CCCCAGAGCGCTCCGCTGCTGTCAFCGT 802
Db 661 ACCTGTGCTGCCCAAGAACATCGTGGCAGCTCAGCCGCGAGGGGCTCCGCTGCTGTCAFCGT 720
QY 803 CTACGTGAACGGTGGGTGGTGTGACG-TGGACCGAGTGTGTCCTGCTG 847
Db 721 CTACGTGGACGGCAGCTGGAGCCCGTGTGGAGCCAGTGTGTCGGGCTG 766

```

```

RESULT 13
BX348193/c
LOCUS
DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.
ACCESSION BX348193
VERSION BX348193.1 GI:30367258
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 843)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3239.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF004ZD01_AF00293_1&cluster=3239.r.

```

Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0BAF0042D01\_AF00293\_1.

## FEATURES

source  
1. .843  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DB008YE02"  
/tissue\_type="NEUROBLASTOMA COT 10-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 24.6%; Score 709.8; DB 13; Length 843;  
Best Local Similarity 95.0%; Pred. No. 4.3e-109;  
Matches 776; Conservative 0; Mismatches 34; Indels 7; Gaps 4;

1694 CAGCTGCACACGCGGAGACGTGAGG--TTGGCCCTAGTGTGCTGTACAGCCCTGCT 1751  
|||||  
824 CACGCGCCAAAGCGGAGAGTGTAGGTTGCCCAAAACCGGCTGTACAGCCCTGTTG 765  
|||||  
1752 AGTCCCATCTGTAGCTGTGACCCCTGGGCTCTGCTCACCAGGCGGAGTATCTGCT 1811  
|||||  
764 AGTCCCATCTGTAGCTGTGAA--CCCTGGGCTCTGCTCACCAGGCGGAGTATCTGCT 706  
|||||  
1812 ATGACCACTGTGGGAGGCCAGCCCTCAGAGCTGTGAGCTGTGCGCTC--AAAAAGCAGTC 1870  
|||||  
705 ATGACCACTGTGGGAGGCCAGCCCTCAGAGCTGTGAGCTGTGCGCTC--AAAAAGCAGTC 646  
|||||  
1871 GTGGAGGAGCTGTGGAGCAGAGTGTGCTGACCTGTGGGAGGAGGCGGCTCCACCT 1930  
|||||  
645 GTGGAGGAGCTGTGG--AGTATGTCTGTGACCTGTGGGAGGAGGCGGCTCCACCT 589  
|||||  
1931 CTACTGTGCTGAGGCGCAGTGTGCTGTAGTCTTACCGAGCAGCTGTGGCGGCTT 1990  
|||||  
588 CTACTGTGCTGAGGCGCAGTGTGCTGTAGTCTTACCGAGCAGCTGTGGCGGCTT 529  
|||||  
1991 TGCCCTGTGGGAGAGCCCTCAGCTGTGGTGTGCGGCAAGCGCTCAAGCTGTCTGTT 2050  
|||||  
528 TGCCCTGTGGGAGAGCCCTCAGCTGTGGTGTGCGGCAAGCGCTCAAGCTGTCTGTT 469  
|||||  
2051 TGCCCGGTGCTGACCTCCCTCAGTACCAATCGGGTCTACTGCTGCTGATGACAC 2110  
|||||  
468 TGCCCGGTGCTGACCTCCCTCAGTACCAATCGGGTCTACTGCTGCTGATGACAC 409  
|||||  
2111 CCAGATGCACTCAAGGAGTGTGCTGAGTGGAGAGCAGCTGTGGGCGGACAGCTGTATCCA 2170  
|||||  
408 CCAGATGCACTCAAGGAGTGTGCTGAGTGGAGAGCAGCTGTGGGCGGACAGCTGTATCCA 349  
|||||  
2171 GGAGCCACGGGTCTGCACTTCAAGGACAGTTACCAAACTGCGCTTATCCATCCACGA 2230  
|||||  
348 GGAGCCACGGGTCTGCACTTCAAGGACAGTTACCAAACTGCGCTTATCCATCCACGA 289  
|||||  
2231 TGTGCCAGCTCCCTGTGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2290  
|||||  
288 TGTGCCAGCTCCCTGTGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 229  
|||||  
2291 TCACATCTGGAATGGCAGCGGCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2350  
|||||  
228 TCACATCTGGAATGGCAGCGGCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 169  
|||||  
2351 CCCAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2410  
|||||  
168 CCCAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 109  
|||||  
2411 GAGCTTCAGCATCACTTCAACATCACCAAGGACACAAAGGTTTCTGCTGCTGCTGCTGCT 2470  
|||||  
108 GAGCTTCAGCATCACTTCAACATCACCAAGGACACAAAGGTTTCTGCTGCTGCTGCTGCT 49

QY 2471 GGAGAGTGAAGCGGGGTCCAGCCCTGTGTGGCCCC 2507  
|||||  
Db 48 GGAGAGTGAAGCGGGGTCCAGCCCTGTGTGGCCCC 12  
|||||

## RESULT 14

LOCUS BX345406  
DEFINITION BX345406 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.

ACCESSION BX345406

VERSION BX345406.1 GI:30369001

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 939)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

COMMENT

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 3529.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0AC023DA05QPI&cluster=3529.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0AC023DA05QPI.

## FEATURES

source

1. .939

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DC023YB10"

/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 24.3%; Score 700.2; DB 13; Length 939;  
Best Local Similarity 94.0%; Pred. No. 1.9e-107;  
Matches 755; Conservative 22; Mismatches 20; Indels 6; Gaps 5;

QY 1 AGCTGGGGCTCCGGGTGAGCGCTAAAGCGGCGCTCCCGCGGCGGCGGCGGCGGCGGCGG 60  
|||||  
Db 142 AGCTGGGGCTCCGGGTGAGCGCTAAAGCGGCGCTCCCGCGGCGGCGGCGGCGGCGGCGG 201  
|||||  
QY 61 GCCCGCGCGCTGCGCGCGCGCGCGCATGGCGCTCCCGCGGCGGCGCTGTGGCCAGCGCTCC 120  
|||||  
Db 202 GCCCGCGCGCTGCGCGCGCGCGCGCATGGCGCTCCCGCGGCGGCGCTGTGGCCAGCGCTCC 261  
|||||  
QY 121 TGGGCATAGTCTGCGCGCTTGGCTCGCGGCTGGGTGCGGCTGCCAGCAGAGTGCCACCGTGG 180  
|||||  
Db 262 TGGGCATAGTCTGCGCGCTTGGCTCGCGGCTGGGTGCGGCTGCCAGCAGAGTGCCACCGTGG 321  
|||||  
QY 181 CCACCCAGTGTGCTGTCACACCCGAGCCTGCTTCCCACTTCTCTGCTGAGGCGGCGGCGG 240  
|||||  
Db 322 CCACCCAGTGTGCTGTCACACCCGAGCCTGCTTCCCACTTCTCTGCTGAGGCGGCGGCGG 381  
|||||  
QY 241 ATGTGTACATCGTCAGACAAAGCAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300  
|||||  
Db 382 ATGTGTACATCGTCAGACAAAGCAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 439  
|||||  
QY 301 AGATCTCTTCAAGTGCACCGGGAGTGGGTGGCGCAGGTGGACCACTGATCGAGCGCA 360  
|||||

Db	440	AGATWTTCTTAAAGTGCAA-GTGGAGTGGGTGGCGGAGGTGACCAAGTATCGAGCGCA	498
Qy	361	GCACAGACGGGAGCAGTGGTGGAGCGGACCAATGAGGTTCGCAATTAATGTCTCAAGGAGC	420
Db	499	-SMCMGACGGGAGCMGTGGGTGGCGCCACCATGGAGTTCGCAATTAATGTCTCAAGGAGC	557
Qy	421	AGGTCGAGAGAGTGTTCGGGCTGGAGGATATCTGTGCCAGTGCCTGGCATGGAGTCTCT	480
Db	558	AGGTCGAGAGAGTGTTCGGGCTGGAGGATATCTGTGCCAGTGCCTGGCATGGAGTCTCT	617
Qy	481	CGGGACACACAGAGTCAAGAGGCTACATCCGATAGCCAGATTGCCAAGAACTTCG	540
Db	618	CGGGACACACAGAGTCAAGAGGCTACATCCGATAGCCATTCGCMAGAMCTTCG	677
Qy	541	AGCAGGAGCGGTGGCCAAAGGAGTGTCCCTGGAGCAGGCGCATGCTGCTGCCCTGCCGTC	600
Db	678	AGCAGGAGCGGTGGCCAAAGGAGTGTCCCTGGAGCAGGCGCATGCTGCTGCCCTGCCGTC	737
Qy	601	CACCGAGGCGATCCCTCAGCGGAGGTGGTCCCGACGAGGACCTGGTGGACC	660
Db	738	CACCGAGGCGATCCCTCAGCGGAGGTGGTCCCGACGAGGACCTGGTGGACC	797
Qy	661	CGTCCCTGGACCCCAATGATATACATCAGCGGAGCACAGCCT-GGTGGTGGCACAGGCC	719
Db	798	CGTCCCTGGACCCCAATGATATACATCAGCGGAGCACAGCCT-GGTGGTGGCACAGGCC	857
Qy	720	CGCTTGTCTGACACGGCCAACTACCTCGTGGCCCAAGAACATCTGGCACCTGCCCGC	779
Db	858	CGCTTGTCTGACACGGCCAACTACCTCGTGGTGGSCAAGMACATCTGGCACCTGCCCGC	917
Qy	780	AGCGCTTCGCTGCTGTCATCGT	802
Db	918	-GGGCTCCGCTGCTGTATCGT	939
RESULT 15			
BX345407			
LOCUS			
DEFINITION			
BX345407 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens			
cDNA clone CS023YB10 5-PRIME, mRNA sequence.			
BX345407			
BX345407.1 GI:30369002			
EST.			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 1040)			
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by Life Technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 3529.f For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0AC023DA05QP2&cluster=3529.f. Contact :			
Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
Faraday Avenue Genoscope sequence ID : CS0AC023DA05QP2.			
FEATURES			
source			
1..1040			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS023YB10"			
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"			
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"			
/note="1st strand cDNA was primed with a NotI-oligo (dT)			
primer. Five prime end enriched, double-strand cDNA was			

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."			
ORIGIN			
Query Match 23.6%; Score 679.4; DB 13; Length 1040;			
Best Local Similarity 94.1%; Pred. No. 6.2e-104;			
Matches 750; Conservative 9; Mismatches 31; Indels 7; Gaps 5;			
Qy	1	AGCTGGGCTCGGAGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGCGCCCGCGCCG	60
Db	161	AGCTGGGCTCGGAGCTGAGCGCTAAAGCGCCCTCCCGCCCGGGGCGCCCGCGCCG	220
Qy	61	GCCCGCCGCTCGGCGCCCGCGCGCATGGCGCTCGCGCCCGGCTGTGGCGCAGCTCC	120
Db	221	GCCCGCCGCTCGGCGCCCGCGCGCATGGCGCTCGCGCCCGGCTGTGGCGCAGCTCC	280
Qy	121	TGGGATAGTCTCGCGCTTGGCTCCCGGCTCGGGTCCCGAGAGTGCACCGTGG	180
Db	281	TGGGATAGTCTCGCGCTTGGCTCCCGGCTCGGGTCCCGAGAGTGCACCGTGG	340
Qy	181	CCAAACCGAGTGGTGGCGCAACCGGACCTGCTTCCCGCTTCTGGTGGAGCCGAG	240
Db	341	CCAAACCGAGTGGTGGCGCAACCGGACCTGCTTCCCGCTTCTGGTGGAGCCGAG	400
Qy	241	ATGTGTACATCGTCAAGAAACAGCCAGTGTGTTGTGTGCAAGCCGTCGCCGACGC	300
Db	401	ATGTGTACATCGTCAAGAAACAGCCAGTGTGTTGTGTGCAAGCCGTCGCCGACGC	460
Qy	301	AGATCTTCTTCAAGTGCACCGGAGTGGTGGCGCCAGGTGAGACACAGTATCGAGCG	360
Db	461	AGATCTTCTTCAAGTGCACCGGAGTGGTGGCGCCAGGTGAGACACAGTATCGAGCG	520
Qy	361	GCACAGCGGAGCAGTGGTGGCGCCAGGTGAGCGCTTCTGGTGGAGCCGAGC	420
Db	521	-SACACCGGAGCAGTGGGCTGCCACCATGAGGTCCGCAATATGTCTCAAGGAGC	579
Qy	421	AGGTCGAGAGGTGTTCGGGCTGGAGGAATATCTGGTCCAGTGGTGGCATGGAGTCT	480
Db	580	AGGTCGAGAGGTGTTCGGGCTGGAGGAATATCTGGTCCAGTGGTGGCATGGAGTCT	639
Qy	481	CGGGCACCAACAGAGTCAAGAGCCCTACATCCGCTATAGCCAGATTGGCAAGAACTTC	540
Db	640	CGGGCACCAACAGAGTCAAGAGCCCTACATCCGCTATAGCCAGATTGGCAAGAACTTC	699
Qy	541	AGCAGGAGCGCTGGCCAAAGGAGTGTCCCTGGAGCAGGCGCATCGTCTGCCCTGCCGTC	600
Db	700	AGCAGGAGCGCTGGCCAAAGGAGTGTCCCTGGAGCAGGCGCATCGTCTGCCCTGCCGTC	759
Qy	601	CACCGAGGCGATCCCTCCAGCGGAGGTGGTGGCTCCCGAAACGAGGACCTGGTGGACC	660
Db	760	CACCGAGGCGATCCCTCCAGCGGAGGTGGTGGCTCCCGAAACGAGGACCTGGTGGACC	819
Qy	661	CGTCCCTGGACCCCAATGTATATACATACCGGGGAGGACAGCTGGTGGTGGAGCGCC	720
Db	820	CGTCCCTGGACCCCAATGTATATACATACCGGGGAGGACAGCTGGTGGTGGAGCGCC	877
Qy	721	GCCTTGTGTACACGGCCAACTACACCTGGTGGCGCCAGAGACATCGTGGCAGCTCGCGCA	780
Db	878	G-CTTGTGTACACGGCAA---CTACACTGGTGGCAAGAACATCTGGCACGTGGCGCA	933
Qy	781	CGCGCTCCGCTGCTGTC	797
Db	934	SCCTBCGCTGTGTCATC	950
Search completed: October 6, 2004, 17:50:24			
Job time : 7199 secs			

**THIS PAGE BLANK (USPTO)**